

Schreiber, David

136375

**From:** Swope, Sheridan  
**Sent:** Saturday, October 23, 2004 8:56 PM  
**To:** STIC-Biotech/ChemLib  
**Cc:** Schreiber, David  
**Subject:** 09/992,095

Please forward to David Schreiber.

David, Would you set up the following search and interference search for 09/992,095?

SID 54: against the NT and AA data bases  
do an oligo search in which the hit sequence can be no longer than the full-length SID 54 (207 AAs).  
i.e. the claim reads a polypeptide consisting of SID 54, or a fragment thereof.

SID 53, residues 1044-1664: against the NT and AA data bases  
do an oligo search in which the hit sequence can be no longer than residues 1044-1664 of SID 53 (621 NTs).  
i.e. the claim reads a polynucleotide consisting of the coding region of SID 53, or a fragment thereof.

Let me know if you have any questions.  
Thank you very much!!  
PS You're the best!!

Sheridan Swope, Ph.D.  
Patent Examiner, AU 1652  
Recombinant Enzymes  
571-272-0943 (voice)  
E02B71 Remsen Bld (Office)  
E02C70 Remsen Bld (Mailbox)

Scan  
this  
page

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:32:03 ; Search time 98 Seconds  
(without alignments)  
584.823 Million cell updates/sec

Title: US-09-992-095B-54  
Perfect score: 207  
Sequence: 1 MHFCGCTLISPEWVLTAAHC.....GVYRVSRFTWIEGVMRNN 207

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1370721 seqs, 324215800 residues

Word size : 0

Total number of hits satisfying chosen parameters: 821245

Minimum DB seq length: 0

Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207	100.0	207	10	US-09-992-600A-54
2	207	100.0	207	10	US-09-924-340-54
3	207	100.0	207	10	US-09-992-095B-54
4	207	100.0	207	10	US-09-999-570-54
5	207	100.0	207	14	US-10-000-489-54
6	207	100.0	207	14	US-10-000-986-54
7	207	100.0	207	14	US-10-154-678-54
8	14	6.8	118	10	US-09-997-003-39
9	13	6.3	18	9	US-09-826-290-247
10	13	6.3	18	15	US-10-264-309-68
11	13	6.3	23	9	US-09-879-792-22
12	13	6.3	23	17	US-10-806-370-22
13	13	6.3	29	9	US-09-826-290-120

14	13	6.3	97	10	US-09-825-751A-8	Sequence 8, Appli
15	13	6.3	149	14	US-10-357-175-20	Sequence 20, Appl
16	13	6.3	149	14	US-10-455-720-20	Sequence 20, Appl
17	13	6.3	157	14	US-10-357-175-23	Sequence 23, Appl
18	13	6.3	157	14	US-10-455-720-23	Sequence 23, Appl
19	13	6.3	191	13	US-10-045-367A-6	Sequence 6, Appli
20	13	6.3	191	14	US-10-170-789-56	Sequence 56, Appl
21	13	6.3	193	14	US-10-106-698-1740	Sequence 4740, Ap
22	13	6.3	199	14	US-10-131-409-98	Sequence 98, Appl
23	13	6.3	199	14	US-10-139-854-98	Sequence 98, Appl
24	13	6.3	199	14	US-10-150-813-98	Sequence 98, Appl
25	13	6.3	199	15	US-10-150-811-98	Sequence 98, Appl
26	12	5.8	12	9	US-09-804-156-33	Sequence 33, Appl
27	12	5.8	12	9	US-09-946-633-18	Sequence 18, Appl
28	12	5.8	12	10	US-09-997-003-53	Sequence 53, Appl
29	12	5.8	12	13	US-10-125-459-18	Sequence 18, Appl
30	12	5.8	12	13	US-10-067-761-33	Sequence 33, Appl
31	12	5.8	12	14	US-10-319-519-33	Sequence 33, Appl
32	12	5.8	159	14	US-10-357-175-24	Sequence 24, Appl
33	12	5.8	159	14	US-10-455-720-24	Sequence 24, Appl
34	11	5.3	12	9	US-09-879-792-16	Sequence 16, Appl
35	11	5.3	12	17	US-10-806-370-16	Sequence 16, Appl
36	11	5.3	30	9	US-09-820-893-111	Sequence 111, App
37	11	5.3	30	15	US-10-607-565-111	Sequence 111, App
38	11	5.3	98	14	US-10-240-730-2	Sequence 2, Appli
39	11	5.3	102	15	US-10-262-511-56	Sequence 56, Appl
40	11	5.3	119	15	US-10-262-511-54	Sequence 54, Appl
41	11	5.3	119	15	US-10-262-511-60	Sequence 60, Appl
42	11	5.3	142	15	US-10-262-511-48	Sequence 48, Appl
43	11	5.3	148	15	US-10-262-511-52	Sequence 52, Appl
44	11	5.3	162	14	US-10-325-745-6	Sequence 6, Appli
45	11	5.3	187	9	US-09-804-156-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-09-992-600A-54  
; Sequence 54, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91 US4 DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-992-600A-54

Query Match 100.0%; Score 207; DB 10; Length 207;  
Best Local Similarity 100.0%; Pred. No. 5,3e-194;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MHFCGCTLISPEWVLTAAHCLEKSPRSSYKYLGAHQEVNLEPHVQIEVSRLEPTR 60

Db 1 MHFCGGTLLSPWVLTAAHCLKSPSPSYKVILGAHQEVNLEPHVQIEVSRLEPFR 60  
Qy 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
Db 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSGQDSGGLVCFEKKYILQGVTSWGLG 180  
Db 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSGQDSGGLVCFEKKYILQGVTSWGLG 180  
Qy 181 CARPNKPGVYVRSRFTWIEGVNRN 207  
Db 181 CARPNKPGVYVRSRFTWIEGVNRN 207

## RESULT 2

US-09-924-340-54  
; Sequence 54, Application US/09924340  
; Publication No. US2003002748A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/924,340  
; CURRENT FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-924-340-54

Query Match 100.0%; Score 207; DB 10; Length 207;  
Best Local Similarity 100.0%; Pred. No. 5.3e-194;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSPSYKVILGAHQEVNLEPHVQIEVSRLEPFR 60  
Db 1 MHFCGGTLLSPWVLTAAHCLKSPSPSYKVILGAHQEVNLEPHVQIEVSRLEPFR 60  
Qy 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
Db 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSGQDSGGLVCFEKKYILQGVTSWGLG 180  
Db 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSGQDSGGLVCFEKKYILQGVTSWGLG 180  
Qy 181 CARPNKPGVYVRSRFTWIEGVNRN 207  
Db 181 CARPNKPGVYVRSRFTWIEGVNRN 207

## RESULT 3

US-09-992-095B-54  
; Sequence 54, Application US/09992095B  
; Publication No. US20030157485A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US5.DIV

; CURRENT APPLICATION NUMBER: US/09/992,095B  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-992-095B-54

Query Match 100.0%; Score 207; DB 10; Length 207;  
Best Local Similarity 100.0%; Pred. No. 5.3e-194;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MHFCGGTLLSPWVLTAAHCLKSPSPSYKVILGAHQEVNLEPHVQIEVSRLEPFR 60  
Db 1 MHFCGGTLLSPWVLTAAHCLKSPSPSYKVILGAHQEVNLEPHVQIEVSRLEPFR 60  
Qy 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
Db 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSGQDSGGLVCFEKKYILQGVTSWGLG 180  
Db 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSGQDSGGLVCFEKKYILQGVTSWGLG 180  
Qy 181 CARPNKPGVYVRSRFTWIEGVNRN 207  
Db 181 CARPNKPGVYVRSRFTWIEGVNRN 207

## RESULT 4

US-09-999-570-54  
; Sequence 54, Application US/09999570  
; Publication No. US20030170628A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: G-091US08DIV  
; CURRENT APPLICATION NUMBER: US/09/999,570  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-999-570-54



Query Match 100.0%; Score 207; DB 10; Length 207;  
Best Local Similarity 100.0%; Pred. No. 5.3e-194; Indels 0; Gaps 0;  
Matches 207; Conservative 0; Mismatches 0;  
QY 1 MHFCGGTLLSPWVLTAAHCLSKSPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
DB 1 MHFCGGTLLSPWVLTAAHCLSKSPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
QY 61 KDIALLLKSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
DB 61 KDIALLLKSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
QY 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180  
DB 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180  
QY 181 CARPNKPGVYVRSRFTVWIEGVMRNN 207  
DB 181 CARPNKPGVYVRSRFTVWIEGVMRNN 207

RESULT 5  
US-10-000-489-54  
; Sequence 54, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US6.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: Jpatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-54

Query Match 100.0%; Score 207; DB 14; Length 207;  
Best Local Similarity 100.0%; Pred. No. 5.3e-194;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MHFCGGTLLSPWVLTAAHCLSKSPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
DB 1 MHFCGGTLLSPWVLTAAHCLSKSPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
QY 61 KDIALLLKSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
DB 61 KDIALLLKSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
QY 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180  
DB 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180  
QY 181 CARPNKPGVYVRSRFTVWIEGVMRNN 207  
DB 181 CARPNKPGVYVRSRFTVWIEGVMRNN 207

RESULT 6  
US-10-000-986-54  
; Sequence 54, Application US/10000986  
; Publication No. US20030096247A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US9.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,986  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: Jpatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-986-54

Query Match 100.0%; Score 207; DB 14; Length 207;  
Best Local Similarity 100.0%; Pred. No. 5.3e-194;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MHFCGGTLLSPWVLTAAHCLSKSPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
DB 1 MHFCGGTLLSPWVLTAAHCLSKSPSSYKVILGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
QY 61 KDIALLLKSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
DB 61 KDIALLLKSSPAVITDKVIPACLPSPNYVADRTCEFTGWTGCTGAGLLKEAQLPV 120  
QY 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180  
DB 121 IENKVCNRYEFLNGRVQSTELCAGHLAGGTDSCQDGGPLVCFEKDKYILQGVTSWGLG 180  
QY 181 CARPNKPGVYVRSRFTVWIEGVMRNN 207  
DB 181 CARPNKPGVYVRSRFTVWIEGVMRNN 207

RESULT 7  
US-10-154-678-54  
; Sequence 54, Application US/10154678  
; Publication No. US20030162186A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 182.US1.REG  
; CURRENT APPLICATION NUMBER: US/10/154,678  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574

```

; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-678-54

Query Match      100.0%; Score 207; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 5.3e-194;
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHFCGTLISEWLVTAACHLEKSPSSYKVIILGAHOEVNLEPHVQIEVSRFLPEPTR 60
Db 1 MHFCGTLISEWLVTAACHLEKSPSSYKVIILGAHOEVNLEPHVQIEVSRFLPEPTR 60

Qy 61 KDIALKLSSPAVITDKVIPACLPSPNVVADRTCEFTTGGETGCTGAGLLKEAQLPV 120
Db 61 KDIALKLSSPAVITDKVIPACLPSPNVVADRTCEFTTGGETGCTGAGLLKEAQLPV 120

Qy 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCGDSGGPLVCFEKKYIILQGVTSWGLG 180
Db 121 IENKVCNRYEFLNGRVQSTELCAGHLAGTDSQCGDSGGPLVCFEKKYIILQGVTSWGLG 180

Qy 181 CARPNKPGVYVRVSRFVTWIEGVRNN 207
Db 181 CARPNKPGVYVRVSRFVTWIEGVRNN 207

```

```

RESULT 8
US-09-997-003-39
; Sequence 39, Application US/09997003
; Publication No. US20030203361A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA003PI
; CURRENT APPLICATION NUMBER: US/09/997,003
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/22157
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/148,680
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-003-39

```

```

Query Match      6.8%; Score 14; DB 10; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 150 TDCQCGDSGGPLVC 163
Db 56 TDCQCGDSGGPLVC 69

RESULT 9
US-09-826-290-247
; Sequence 247, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
; APPLICANT: Durham, L. Kathryn
; APPLICANT: Friedman, David L.
; APPLICANT: Herath, Herath Mudiyanselage Athula Chandrasiri
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Parekh, Rajesh Bhikhu

```

```

; APPLICANT: Potter, David M.
; APPLICANT: Rohlf, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
; APPLICANT: Williams, Stephen A.
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 247
; LENGTH: 18
; TYPE: PRT
; ORGANISM: homo sapien
US-09-826-290-247

```

```

Query Match      6.3%; Score 13; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 151 DSCQCGDSGGPLVC 163
Db 1 DSCQCGDSGGPLVC 13

```

```

RESULT 10
US-10-264-309-68
; Sequence 68, Application US/10264309
; Publication No. US20040022794A1
; GENERAL INFORMATION:
; APPLICANT: DURHAM, L. KATHRYN
; APPLICANT: FRIEDMAN, DAVID L.
; APPLICANT: HERATH, HERATH
; APPLICANT: KIMMEL, LIDA H.
; APPLICANT: PAREKH, RAJESH B.
; APPLICANT: POTTER, DAVID M.
; APPLICANT: ROHLFF, CHRISTIAN
; APPLICANT: SILBER, B. MICHAEL
; APPLICANT: SNYDER, PETER J.
; APPLICANT: SOARES, HOLLY D.
; APPLICANT: STIGER, THOMAS R.
; APPLICANT: SUNDERLAND, P. TREY
; APPLICANT: TOWNSEND, ROBERT R.
; APPLICANT: WHITE, W. FROST
; APPLICANT: WILLIAMS, STEPHEN A.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, POLYPEPTIDES AND USES THEREFOR,
; TITLE OF INVENTION: INCLUDING DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE
; FILE REFERENCE: POA-002.01
; CURRENT APPLICATION NUMBER: US/10/264,309
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: 60/326,708
; PRIOR FILING DATE: 2001-10-03
; NUMBER OF SEQ ID NOS: 491
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 68
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-309-68

```

```

Query Match      6.3%; Score 13; DB 15; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
; FEATURE:
; OTHER INFORMATION: BLOCKS BL00134B
US-10-806-370-22

Query Match          6.3%; Score 13; DB 17; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-05; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0;

QY 151 DSCQDSDGGPLVC 163
Db 1 DSCQDSDGGPLVC 13

RESULT 11
US-09-879-792-22
; Sequence 22, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; TITLE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12 (Docket No. US20020061850A1 LIO-81-WO)
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: BLOCKS BL00134B
US-09-879-792-22

Query Match          6.3%; Score 13; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.2e-05; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0;

QY 151 DSCQDSDGGPLVC 163
Db 1 DSCQDSDGGPLVC 13

RESULT 12
US-10-806-370-22
; Sequence 22, Application US/10806370
; Publication No. US20040209327A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; TITLE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/10/806,370
; CURRENT FILING DATE: 2004-03-23
; PRIOR APPLICATION NUMBER: US/09/879,792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12 (Docket No. LIO-81-WO)
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: BLOCKS BL00134B
US-09-826-290-120

Query Match          6.3%; Score 13; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.1e-05; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0;

QY 151 DSCQDSDGGPLVC 163
Db 12 DSCQDSDGGPLVC 24

RESULT 14
US-09-825-751A-8
; Sequence 8, Application US/09825751A
; Publication No. US20030065140A1
; GENERAL INFORMATION:
; APPLICANT: CuraGen Corporation
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Herrman, John L.
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-750
```

```
; CURRENT APPLICATION NUMBER: US/09/825,751A
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/194,314
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 60/225,693
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 8
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-825-751A-8
```

```
Query Match      6.3%; Score 13; DB 10; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 151 DSCQDGGGGLVC 163
   |||||
Db 41 DSCQDGGGGLVC 53
```

```
RESULT 15
US-10-357-175-20
; Sequence 20, Application US/10357175
; Publication No. US20030170707A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
; FILE REFERENCE: D6192CIP/D/CIP
; CURRENT APPLICATION NUMBER: US/10/357,175
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/650,371
; PRIOR FILING DATE: 2000-08-28
; NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 20
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; OTHER INFORMATION: protease domain of protease M (Prom)
US-10-357-175-20
```

```
Query Match      6.3%; Score 13; DB 14; Length 149;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 151 DSCQDGGGGLVC 163
   |||||
Db 137 DSCQDGGGGLVC 149
```

```
Search completed: October 27, 2004, 09:40:11
Job time : 98 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: October 27, 2004, 09:12:42 ; Search time 107 Seconds

(without alignments)  
693.990 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 207

Sequence: 1 MHFGCTGLSPENLTAHC.....GVVVRVSRFTWIEGVMRNN 207

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1396543

Minimum DB seq length: 0

Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207	100.0	207	6	ABR48479 Human Pla
2	26	12.6	47	7	ADH34460 Plasmin,
3	18	8.7	95	4	ABG01924 Novel hum
4	17	8.2	84	5	ABP02684 Human ORF
5	14	6.8	118	4	ABW71662 Human col
6	13	6.3	113	8	ADO71577 Amino aci
7	13	6.3	14	5	AAE17240 Human tra
8	13	6.3	14	5	AAE18999 Human mat
9	13	6.3	18	4	ABBS2232 Human API
10	13	6.3	18	4	ABBS2105 Human API
11	13	6.3	18	6	ABR58948 Alzheimer
12	13	6.3	18	8	ADN31742 Human Alz
13	13	6.3	24	5	AAE17246 Serine pr
14	13	6.3	24	5	AAE19007 Human mat
15	13	6.3	35	5	AAE17239 Human tra
16	13	6.3	35	5	AAE19008 Human tra
17	13	6.3	97	4	AAW50210 Human pla
18	13	6.3	141	7	ADG75740 Human pro
19	13	6.3	149	4	AAU68926 Human pro
20	13	6.3	149	7	ADE13246 Protease
21	13	6.3	149	8	ADH78455 Human pro
22	13	6.3	156	6	ABR41514 Human DIT
23	13	6.3	157	4	AAU68929 Human pro
24	13	6.3	157	7	ADE13374 Protease
25	13	6.3	157	8	ADH78458 Human TAD

26	13	6.3	181	8	ADI16332 Human pro
27	13	6.3	191	6	ABG75787 Serine pr
28	13	6.3	193	4	AAG73966 Human col
29	13	6.3	194	8	ADM79155 Human del
30	13	6.3	199	7	ADG42737 Human bet
31	13	6.3	199	7	ADJ55806 Peptide h
32	13	6.3	199	8	ADM76644 Human NOV
33	12	5.8	12	3	AAV72111 Peptide f
34	12	5.8	12	4	AAV72111 Peptide f
35	12	5.8	12	5	ABG30800 Human ser
36	12	5.8	12	5	AAE17931 Human gen
37	12	5.8	159	4	AAU68930 Human pro
38	12	5.8	159	7	ADE13375 Protease
39	12	5.8	159	8	ADH78459 Human TMP
40	11	5.3	11	8	ADO71578 Amino aci
41	11	5.3	16	3	AAE23914 Porcine v
42	11	5.3	30	3	AAE08953 Human sec
43	11	5.3	70	5	AAU76373 Human Mar
44	11	5.3	84	2	AAW01208 Serine pr
45	11	5.3	84	4	AAW50605 Flea seri

## ALIGNMENTS

RESULT 1  
ABR48479  
ID ABR48479 standard; protein; 207 AA.  
XX  
AC ABR48479;  
DT 13-JUN-2003 (first entry)  
XX  
DE Human Plasminute.  
XX  
KW Human; GENSET; therapeutic; therapy.  
XX  
OS Homo sapiens.  
XX  
FN WO200294864-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 06-AUG-2001; 2001WO-IB001715.  
XX  
PR 25-MAY-2001; 2001US-0293574P.  
PR 15-JUN-2001; 2001US-0298698P.  
PR 29-JUN-2001; 2001US-0302277P.  
PR 13-JUL-2001; 2001US-0305456P.  
XX  
PA (GEST ) GENSET.  
XX  
PI Bejanin S, Tanaka H;  
DR WPI; 2003-129412/12.  
DR N-PSDB; ACC51086.  
XX  
PS New GENSET polynucleotides and polypeptides, useful for preparing a  
composition for treating GENSET-related disorders and as reagents in  
assays to quantitatively determined levels of GENSET expression in  
biological samples.  
XX  
PS Claim 2; Page 447-448; 505pp; English.

The present invention relates to novel human GENSET coding sequences (ACC51060-ACC51115) and proteins (ABR48453-ABR48508). The GENSET sequences are useful for preparing a composition for treating GENSET-related disorders. They can also be used as markers for tissues in which the corresponding protein is preferentially expressed, as molecular weight markers on Southern gels, as chromosome markers or tags to identify chromosomes, and as reagents in assays to quantitatively determined levels of GENSET expression in biological samples

SQ Sequence 207 AA;  
Query Match 100.0%; Score 207; DB 6; Length 207;  
Best Local Similarity 100.0%; Pred. No. 7e-95;  
Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MHFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
DB 1 MHFCGGTLLSPWVLTAAHCKLEKSPSSYKVLGAHQEVNLEPHVQIEVSRFLFLEPTR 60  
QY 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTGTCFITGNETGCTGAGLLKEAQLPV 120  
DB 61 KDIALLLSSPAVITDKVIPACLPSPNYVADRTGTCFITGNETGCTGAGLLKEAQLPV 120  
QY 121 TENKVCNRYEFLNGRVQSTELCAGHLAGTSCQDGGPLVCFPEKDKYIILQGVTSWGLG 180  
DB 121 TENKVCNRYEFLNGRVQSTELCAGHLAGTSCQDGGPLVCFPEKDKYIILQGVTSWGLG 180  
QY 181 CARPNKPGVYVRSFVTVWIEGVRNN 207  
DB 181 CARPNKPGVYVRSFVTVWIEGVRNN 207  
RESULT 2  
ADH34460  
ID ADH34460 standard; protein; 47 AA.  
AC ADH34460;  
XX  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Plasmin, a megin ligand useful for treating renal failure.  
XX  
KW Plasmin; megin ligand; serpin; protease inhibitor;  
KW glomerular mesangial cell; drug screening;  
KW glomerular mesangial inflammation; renal failure.  
XX  
OS Unidentified.  
XX  
XX WO2003066089-A1.  
XX  
PD 14-AUG-2003.  
XX  
PF 07-FEB-2003; 2003WO-JP001316.  
XX  
PR 08-FEB-2002; 2002JP-00033164.  
XX  
PA (KURO/) KUROKAWA K.  
PA (MIYA/) MIYATA T.  
PI Kurokawa K, Miyata T;  
XX  
XX WPI; 2003-646264/61.  
XX  
PT Megin ligand containing plasmin and/or trypsin for treatment of  
PT glomerular mesangial inflammation.  
XX  
PS Disclosure; Page 12; 63pp; Japanese.  
XX  
XX The invention relates to megin ligands comprising plasmin and/or trypsin  
CC as the active component. Megin is a member of the serpin family of  
CC protease inhibitors, binding to the serine proteases plasmin and trypsin,  
CC thereby inhibiting their activity. Megin is expressed in glomerular  
CC mesangial cells and its overexpression is a cause of renal failure. The  
CC invention also relates to a medical composition for treatment and  
CC prevention of glomerular mesangial inflammation; a method of screening  
CC for compounds that inhibit binding of megin to megin ligands; and the  
CC compounds thus identified. The megin ligands can be used in the  
CC treatment and prevention of glomerular mesangial inflammation and renal  
CC failure. The present sequence represents plasmin, a megin ligand that  
CC may be used in the invention.  
XX  
SQ Sequence 47 AA;

Query Match 12.6%; Score 26; DB 7; Length 47;  
Best Local Similarity 100.0%; Pred. No. 4e-06;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MHFCGGTLLSPWVLTAAHCKLEKSPR 26  
DB 1 MHFCGGTLLSPWVLTAAHCKLEKSPR 26  
RESULT 3  
ABG01924  
ID ABG01924 standard; protein; 95 AA.  
XX  
AC ABG01924;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #1915.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS66111.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity.  
XX  
XX Claim 20; SEQ ID NO 32283; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
sequences. (I) is useful as hybridisation probes, polymerase chain  
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
and in recombinant production of (II). The polynucleotides are also used  
in diagnostics as expressed sequence tags for identifying expressed  
genes. (I) is useful in gene therapy techniques to restore normal  
activity of (II) or to treat disease states involving (II). (II) is  
useful for generating antibodies against it, detecting or quantitating a  
polypeptide in tissue, as molecular weight markers and as a food  
supplement. (II) and its binding partners are useful in medical imaging  
of sites expressing (II). (I) and (II) are useful for treating disorders  
involving aberrant protein expression or biological activity. The  
polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
amino acid sequences of the invention. Note: The sequence data for this  
patent did not appear in the printed specification, but was obtained in  
electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 95 AA;

Query Match 8.7%; Score 18; DB 4; Length 95;



XX microplasmin; loop domain; alpha2-antiplasmin; factor D; blood clot;  
 KW thrombotic disorder; peripheral arterial occlusion;  
 KW deep venous thrombosis; occlusion; blood vessel.  
 XX Unidentified.  
 OS  
 XX WO2004045558-A2.  
 PN  
 XX 03-JUN-2004.  
 PD  
 XX 18-NOV-2003; 2003WO-US037221.  
 XX  
 PF 18-NOV-2002; 2002US-0427152P.  
 XX  
 PR (HARD ) HARVARD COLLEGE.  
 XX  
 PA Reed GL;  
 XX  
 PI  
 XX WPI; 2004-420534/39.  
 DR  
 XX Novel alpha2-antiplasmin resistant microplasmin polypeptide comprising  
 PT heterologous loop domain sequence, useful for treating thrombotic  
 PT disorders such as deep venous thrombosis.  
 XX  
 XX Disclosure; Page 2; 30pp; English.  
 PS  
 XX The specification describes a microplasmin polypeptide, comprising a  
 CC heterologous loop domain sequence, where the polypeptide is resistant to  
 CC alpha2-antiplasmin inhibition compared to a wild-type microplasmin. The  
 CC heterologous loop domain comprises at least 4 or 10 consecutive amino  
 CC acids of a factor D loop domain, and is present in microplasmin loop 3,  
 CC loop 5, loop 6 or loop 7. The microplasmin polypeptide of the invention  
 CC is useful for dissolving a blood clot. It is useful for treating a  
 CC subject suffering from or at risk of developing a thrombotic disorder,  
 CC such as peripheral arterial occlusion, deep venous thrombosis or other  
 CC disorder associated with occlusion of a blood vessel. AD071576-AD071579  
 CC represent peptides from microplasmin loop 3, loop 5, loop 6 and loop 7,  
 CC respectively. These peptides are replaced with factor D loop domain  
 CC peptides to create microplasmin polypeptides of the invention.  
 XX  
 XX Sequence 13 AA;  
 SQ  
 Query Match 6.3%; Score 13; DB 8; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 5.1;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 18 AHCLEKSPRPSSY 30  
 DB 1 AHCLEKSPRPSSY 13  
 |||||  
 RESULT 7  
 AAEL7240  
 ID AAEL7240 standard; peptide; 14 AA.  
 XX  
 AC AAEL7240;  
 XX  
 DT 18-APR-2002 (first entry)  
 XX  
 DE Human transmembrane serine protease-related type I fibronectin domain #1.  
 XX  
 KW Transmembrane serine protease; gene therapy; metastasis; tumour;  
 KW chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;  
 KW atherosclerosis; neurodegenerative disease; neuroprotective; cystostatic;  
 KW pathogenic infection; antiinflammatory; antiarteriosclerotic;  
 KW antibacterial; type I fibronectin domain.  
 XX  
 OS Unidentified.  
 OS  
 XX WO200196538-A2.  
 PN  
 XX 20-DEC-2001.  
 PD

XX 12-JUN-2001; 2001WO-EP006618.  
 PF  
 XX 13-JUN-2000; 2000US-0211224P.  
 PR  
 PR 13-APR-2001; 2001US-0283353P.  
 PR 16-APR-2001; 2001US-0283648P.  
 XX  
 XX (FARB ) BAYER AG.  
 DA  
 XX Xiao Y, Gedrich R;  
 PI  
 XX WPI; 2002-098065/13.  
 DR  
 XX Novel isolated polynucleotide encoding transmembrane serine protease  
 PT polypeptide, for treating chronic obstructive pulmonary disease, tumor  
 PT angiogenesis, inflammation, atherosclerosis and neurodegenerative  
 PT disease.  
 XX  
 XX Disclosure; Fig 3; 120pp; English.  
 PS  
 XX The present invention relates to an isolated polynucleotide encoding a  
 CC transmembrane serine protease polypeptide. Transmembrane serine protease  
 CC gene is useful in gene therapy. The invention also relates to a  
 CC pharmaceutical composition which is useful for modulating the activity of  
 CC transmembrane serine protease in a disease, such as chronic obstructive  
 CC pulmonary disease (COPD), metastasis of malignant cells, tumour  
 CC angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease  
 CC or pathogenic infection. Transmembrane serine protease is useful as a  
 CC bait protein in a two-hybrid or three-hybrid assay. The polypeptide is  
 CC useful for generating antibodies against it and in various assay systems.  
 CC The present sequence is a human transmembrane serine protease-related  
 CC type I fibronectin domain  
 XX  
 XX Sequence 14 AA;  
 SQ  
 Query Match 6.3%; Score 13; DB 5; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 5.3;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 151 DSCQDGGGGLVC 163  
 DB 2 DSCQDGGGGLVC 14  
 |||||  
 RESULT 8  
 AAEL18999  
 ID AAEL18999 standard; peptide; 14 AA.  
 XX  
 AC AAEL18999;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Human matritase-like serine protease related fibronectin domain #1.  
 XX  
 KW Human; matritase-like serine protease; cardiovascular; antiinflammatory;  
 KW cystostatic; chronic obstructive pulmonary disease; pulmonary; vulnery;  
 KW viricide; nervous system disorder; extracellular matrix degradation;  
 KW antimicrobial; neuroprotective; cancer; cardiovascular disorder; wound;  
 KW adenocarcinoma; neurodegenerative disease; Alzheimer's disease; melanoma;  
 KW Genstmann-Straussler Syndrome; Creutzfeldt-Jakob disease; scrapie;  
 KW Parkinson's disease; enzyme; type I fibronectin domain.  
 XX  
 OS Unidentified.  
 OS  
 XX WO200208392-A2.  
 PN  
 XX 31-JAN-2002.  
 PD  
 XX 16-JUL-2001; 2001WO-EP008182.  
 PF  
 XX 25-JUL-2000; 2000US-0220807P.  
 PR  
 PR 02-APR-2001; 2001US-0280109P.  
 XX



PA (FARB ) BAYER AG.  
XX  
PI Xiao Y;  
XX  
XX WPI; 2002-195871/25.  
XX  
XX New matrilysin-like serine protease proteins and polynucleotides, useful  
PT for treating matrilysin-like serine protease dysfunction related  
PT diseases, e.g. cancer, pulmonary disease, wounds, inflammation or viral  
PT infections.  
XX  
XX Disclosure; Fig 9; 140pp; English.  
XX  
XX The patent discloses novel matrilysin-like serine protease proteins and  
CC their corresponding polynucleotides. The invention further relates to  
CC reagents and methods of regulating human matrilysin-like serine protease  
CC activity. Pharmaceutical compositions comprising the reagents of the  
CC invention are useful for modulating the activity of a matrilysin-like  
CC serine protease in a disease. The reagents are also useful for treating  
CC or ameliorating matrilysin-like serine protease dysfunction related  
CC diseases such as cancer, chronic obstructive pulmonary disease, central  
CC or peripheral nervous system disorder and cardiovascular disorder. The  
CC human matrilysin-like serine protease gene provides a therapeutic target  
CC to decrease extracellular matrix degradation, in particular for treating  
CC or preventing metastatic cancer, e.g. adenocarcinoma, melanoma, cancers  
CC of the adrenal gland, bladder, bone breast, cervix, gall bladder, liver,  
CC lung, ovary, pancreas, prostate, testis or uterus. Sequences of the  
CC invention are also useful for treating neurodegenerative diseases (e.g.,  
CC Alzheimer's disease, Parkinson's disease), wounds, inflammation or viral  
CC infections. Matrilysin-like serine protease activity can be used to  
CC degrade, prion protein amyloid plaques of Genstmann-Straussler Syndrome,  
CC Creutzfeldt-Jakob disease and Scrapie. The present sequence is human  
CC matrilysin-like serine protease related type I fibronectin domain  
XX  
SQ Sequence 14 AA;  
Query Match 6.3%; Score 13; DB 5; Length 14;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 DSCQGDSDGGPLVC 163  
DB 2 DSCQGDSDGGPLVC 14  
RESULT 9  
ABB52232  
ID ABB52232 standard; peptide; 18 AA.  
XX  
AC ABB52232;  
XX  
XX 08-FEB-2002 (first entry)  
XX  
XX Human API-180 tryptic digest peptide #4.  
XX  
XX Human; neuroprotective; nootropic; gene therapy; vaccine;  
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;  
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;  
KW Expression Reference Protein Isoform; ERPI; proteolysis.  
XX  
OS Homo sapiens.  
XX  
XX WO200175454-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 03-APR-2001; 2001WO-US010908.  
XX  
XX 03-APR-2000; 2000US-0194504P.  
XX  
XX 28-NOV-2000; 2000US-0253647P.  
XX  
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX  
XX (PFIZ ) PFIZER INC.  
XX  
XX 11-OCT-2001.  
XX  
XX 03-APR-2001; 2001WO-US010908.  
XX  
XX 03-APR-2000; 2000US-0194504P.  
XX  
XX 28-NOV-2000; 2000US-0253647P.  
XX  
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX  
XX (PFIZ ) PFIZER INC.  
XX  
XX Durham KL, Friedman DL, Herath HM, Kimmel LH, Parekh RB;  
PI Potter DM, Rohlf C, Silber BM, Stiger TR, Sunderland PT;  
PI Townsend RR, White F, Williams SA;  
XX  
XX WPI; 2001-639384/73.  
XX  
XX Screening for Alzheimer's disease in a mammal, by making two-dimensional  
PT array of a feature whose relative abundance correlates with disease, and  
PT comparing with abundance of the feature in samples of healthy persons.  
XX  
XX Example; Page 31; 162pp; English.  
XX  
XX The invention relates to methods for the screening, diagnosis and  
CC prognosis of Alzheimer's disease. The methods involve the detection of  
CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-  
CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or  
CC plasma. The abundance of the AFs and APIs is then normalized to an  
CC Expression Reference Protein Isoform (ERPI) in order to determine whether  
CC a patient is suffering from, or has a predisposition to, Alzheimer's  
CC Disease. The relative abundance of the AFs and APIs correlates with the  
CC severity of Alzheimer's Disease. The present sequence is a peptide  
CC produced from an API by proteolysis  
XX  
SQ Sequence 18 AA;  
Query Match 6.3%; Score 13; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 DSCQGDSDGGPLVC 163  
DB 1 DSCQGDSDGGPLVC 13  
RESULT 10  
ABB52105  
ID ABB52105 standard; peptide; 18 AA.  
XX  
XX ABB52105;  
XX  
XX 08-FEB-2002 (first entry)  
XX  
XX Human API-26 tryptic digest peptide #3.  
XX  
XX Human; neuroprotective; nootropic; gene therapy; vaccine;  
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;  
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;  
KW Expression Reference Protein Isoform; ERPI; proteolysis.  
XX  
XX Homo sapiens.  
XX  
XX WO200175454-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 03-APR-2001; 2001WO-US010908.  
XX  
XX 03-APR-2000; 2000US-0194504P.  
XX  
XX 28-NOV-2000; 2000US-0253647P.  
XX  
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX  
XX (PFIZ ) PFIZER INC.  
XX  
XX Durham KL, Friedman DL, Herath HM, Kimmel LH, Parekh RB;  
PI Potter DM, Rohlf C, Silber BM, Stiger TR, Sunderland PT;  
PI Townsend RR, White F, Williams SA;  
XX  
XX WPI; 2001-639384/73.  
XX  
XX Screening for Alzheimer's disease in a mammal, by making two-dimensional  
PT array of a feature whose relative abundance correlates with disease, and  
PT comparing with abundance of the feature in samples of healthy persons.  
XX  
XX Example; Page 31; 162pp; English.  
XX  
XX The invention relates to methods for the screening, diagnosis and  
CC prognosis of Alzheimer's disease. The methods involve the detection of  
CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-  
CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or  
CC plasma. The abundance of the AFs and APIs is then normalized to an  
CC Expression Reference Protein Isoform (ERPI) in order to determine whether  
CC a patient is suffering from, or has a predisposition to, Alzheimer's  
CC Disease. The relative abundance of the AFs and APIs correlates with the  
CC severity of Alzheimer's Disease. The present sequence is a peptide  
CC produced from an API by proteolysis  
XX  
SQ Sequence 18 AA;  
Query Match 6.3%; Score 13; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 DSCQGDSDGGPLVC 163  
DB 1 DSCQGDSDGGPLVC 13

XX PS Example; Page 28; 162pp; English.

XX CC The invention relates to methods for the screening, diagnosis and prognosis of Alzheimer's disease. The methods involve the detection of Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or plasma. The abundance of the AFs and APIs is then normalised to an CC Expression Reference Protein Isoform (ERPI) in order to determine whether CC a patient is suffering from, or has a predisposition to, Alzheimer's CC Disease. The relative abundance of the AFs and APIs correlates with the CC severity of Alzheimer's Disease. The present sequence is a peptide CC produced from an API by proteolysis

XX SQ Sequence 18 AA;

Query Match 6.3%; Score 13; DB 4; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDGGGLVLC 163  
 Db 1 DSCQDGGGLVLC 13  
 |||||

RESULT 11

ABR58948  
 ID ABR58948 standard; peptide; 18 AA.  
 AC ABR58948;  
 DT 11-JUL-2003 (first entry)  
 DE Alzheimer's Disease-associated protein isoform, API-26, SEQ ID 68.  
 KW Nootropic; Neuroprotective; Alzheimer's disease; API; human;  
 KW Alzheimer's Disease-associated protein isoform.  
 XX Homo sapiens.  
 XX WO2003028543-A2.  
 XX 10-APR-2003.  
 XX 03-OCT-2002; 2002WO-US031642.  
 XX 03-OCT-2001; 2001US-0326708P.  
 XX (PFIZ ) PFIZER PROD INC.  
 XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.

Durham LK, Friedman DL, Herath HMCAC, Kimmel LH, Parekh RB;  
 PI Potter DM, Rohlf C, Silber BM, Snyder PJ, Soares HD, Stiger TR;  
 PI Sunderland PT, Townsend RR, White WF, Williams SA;  
 XX WPI; 2003-371957/35.

XX Screening or diagnosing of Alzheimer's disease (AD) determine the stage  
 PT or severity of AD in a subject, comprises analyzing a test sample of body  
 PT fluid from the subject by 2-dimensional electrophoresis.

XX Claim 2; Page 41; 179pp; English.

XX The present invention relates to methods for screening or diagnosing  
 CC Alzheimer's disease (AD) to determine the stage or severity of AD in a  
 CC subject, to identify subject at risk of developing AD, or to monitor the  
 CC effect of therapy administered. The methods comprise analysing a test  
 CC sample of body fluid by 2-dimensional electrophoresis to generate a 2-  
 CC dimensional array of AD-associated features (AFs). The method  
 CC alternatively comprises quantitatively detecting in a sample of body  
 CC fluid from the subject, one or more AD-associated protein isoforms (APIs;  
 CC ABR58710-ABR59184)

XX SQ Sequence 18 AA;

Query Match 6.3%; Score 13; DB 6; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDGGGLVLC 163  
 Db 1 DSCQDGGGLVLC 13  
 |||||

RESULT 12

ADN31742  
 ID ADN31742 standard; peptide; 18 AA.  
 AC ADN31742;  
 DT 01-JUL-2004 (first entry)  
 DE Human Alzheimer's disease-API tryptic digest peptide - SEQ ID 68.  
 KW Alzheimer's disease; nootropic; neuroprotective; cerebrospinal fluid;  
 KW CSF; Alzheimer's disease-associated protein isoform; API; tryptic digest;  
 XX human.  
 XX Homo sapiens.  
 XX EPI408333-A2.  
 XX 14-APR-2004.  
 XX 03-OCT-2002; 2002EP-00256893.  
 XX 03-OCT-2002; 2002US-0326708P.  
 XX (PFIZ ) PFIZER PROD INC.  
 XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.

Durham LK, Friedman DL, Herath HMCAC, Kimmel LH, Parekh RB;  
 PI Potter DM, Rohlf C, Silber BM, Snyder PJ, Soares HD, Stiger TR;  
 PI Sunderland PT, Townsend RR, White WF, Williams SA;  
 XX WPI; 2004-318939/30.

XX Screening or diagnosis of Alzheimer's disease (AD) in subject,  
 PT determining stage or severity of AD, identifying subject at risk of  
 PT developing AD, or monitoring effect of therapy, by detecting Alzheimer's  
 PT disease-Associated Features.

XX Example; SEQ ID NO 68; 208pp; English.

XX The invention relates to a novel method for screening or diagnosis of  
 CC Alzheimer's disease (AD) in a subject, determining the stage or severity  
 CC of AD, identifying a subject at risk of developing AD or monitoring the  
 CC effect of therapy administered to a subject having AD, by analysing body  
 CC fluid to generate a two-dimensional array of Alzheimer's disease-  
 CC associated features (AFs) such as AF-200, AF-201, AF-202, AF-203, AF-204,  
 CC AF-205, etc., and comparing the abundance of AFs with a control. The  
 CC method of the invention has nootropic and neuroprotective applications  
 CC and may be useful for screening or diagnosis of Alzheimer's disease (AD)  
 CC in a subject, determining the stage or severity of AD in a subject,  
 CC identifying a subject at risk of developing AD or monitoring the effect  
 CC of therapy administered to a subject having AD. The body fluid is  
 CC cerebrospinal fluid (CSF). The current sequence is that of a human  
 CC Alzheimer's disease-associated protein isoform (API) tryptic digest  
 CC peptide of the invention.

XX SQ Sequence 18 AA;

Query Match 6.3%; Score 13; DB 8; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 6;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPLVC 163  
 DB 1 DSCQDSDGGPLVC 13

## RESULT 13

AAE17246  
 ID AAE17246 standard; peptide; 24 AA.

XX AAE17246;

XX 18-APR-2002 (first entry)

DE Serine protease peptide #2 of the trypsin family.

XX Transmembrane serine protease; gene therapy; metastasis; tumour;  
 KW chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;  
 KW atherosclerosis; neurodegenerative disease; neuroprotective; cytostatic;  
 KW pathogenic infection; antiinflammatory; antiarteriosclerotic;  
 KW antibacterial.

XX Unidentified.

OS WO200196538-A2.

XX 20-DEC-2001.

XX 12-JUN-2001; 2001WO-EP006618.

XX 13-JUN-2000; 2000US-0211224P.

PR 13-APR-2001; 2001US-0283353P.

PR 16-APR-2001; 2001US-0283648P.

XX (FARB ) BAYER AG.

XX Xiao Y, Gedrich R;

XX WPI; 2002-098065/13.

XX Novel isolated polynucleotide encoding transmembrane serine protease  
 PT polypeptide, for treating chronic obstructive pulmonary disease, tumor  
 PT angiogenesis, inflammation, atherosclerosis and neurodegenerative  
 PT disease.

XX Disclosure; Fig 3; 120pp; English.

XX The present invention relates to an isolated polynucleotide encoding a  
 CC transmembrane serine protease polypeptide. Transmembrane serine protease  
 CC gene is useful in gene therapy. The invention also relates to a  
 CC pharmaceutical composition which is useful for modulating the activity of  
 CC transmembrane serine protease in a disease, such as chronic obstructive  
 CC pulmonary disease (COPD), metastasis of malignant cells, tumour  
 CC angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease  
 CC or pathogenic infection. Transmembrane serine protease is useful as a  
 CC bait protein in a two-hybrid or three-hybrid assay. The polypeptide is  
 CC useful for generating antibodies against it and in various assay systems.  
 CC The present sequence is a serine protease peptide of the trypsin family

XX Sequence 24 AA;

Query Match 6.3%; Score 13; DB 5; Length 24;

Best Local Similarity 100.0%; Pred. No. 7; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPLVC 163

DB 1 DSCQDSDGGPLVC 13

## RESULT 14

AAE19007  
 ID AAE19007 standard; peptide; 24 AA.

XX

AC AAE19007;

XX 21-MAY-2002 (first entry)

XX Human matrilysin-like serine protease related peptide #3.

XX Human; matrilysin-like serine protease; cardiovascular; antiinflammatory;  
 KW cytostatic; chronic obstructive pulmonary disease; pulmonary; vulnary;  
 KW viricide; nervous system disorder; extracellular matrix degradation;  
 KW antimicrobial; neuroprotective; cancer; cardiovascular disorder; wound;  
 KW adenocarcinoma; neurodegenerative disease; Alzheimer's disease; melanoma;  
 KW Genstmann-Straussler Syndrome; Creutzfeldt-Jakob disease; scrapie;  
 KW Parkinson's disease; enzyme.

XX Unidentified.

XX WO200208392-A2.

XX 31-JAN-2002.

XX 16-JUL-2001; 2001WO-EP008182.

XX 25-JUL-2000; 2000US-0220807P.

XX 02-APR-2001; 2001US-0280109P.

XX (FARB ) BAYER AG.

XX Xiao Y;

XX WPI; 2002-195871/25.

XX New matrilysin-like serine protease proteins and polynucleotides, useful  
 PT for treating matrilysin-like serine protease dysfunction related  
 PT diseases, e.g. cancer, pulmonary disease, wounds, inflammation or viral  
 PT infections.

XX Disclosure; Fig 9; 140pp; English.

XX The patent discloses novel matrilysin-like serine protease proteins and  
 CC their corresponding polynucleotides. The invention further relates to  
 CC reagents and methods of regulating human matrilysin-like serine protease  
 CC activity. Pharmaceutical compositions comprising the reagents of the  
 CC invention are useful for modulating the activity of a matrilysin-like  
 CC serine protease in a disease. The reagents are also useful for treating  
 CC or ameliorating matrilysin-like serine protease dysfunction related  
 CC diseases such as cancer, chronic obstructive pulmonary disease, central  
 CC or peripheral nervous system disorder and cardiovascular disorder. The  
 CC human matrilysin-like serine protease gene provides a therapeutic target  
 CC to decrease extracellular matrix degradation, in particular for treating  
 CC or preventing metastatic cancer, e.g. adenocarcinoma, melanoma, cancers  
 CC of the adrenal gland, bladder, bone breast, cervix, gall bladder, liver,  
 CC lung, ovary, pancreas, prostate, testis or uterus. Sequences of the  
 CC invention are also useful for treating neurodegenerative diseases (e.g.,  
 CC Alzheimer's disease, Parkinson's disease), wounds, inflammation or viral  
 CC infections. Matrilysin-like serine protease activity can be used to  
 CC degrade, prion protein amyloid plaques of Genstmann-Straussler Syndrome,  
 CC Creutzfeldt-Jakob disease and Scrapie. The present sequence is human  
 CC matrilysin-like serine protease related peptide of trypsin family

XX Sequence 24 AA;

Query Match 6.3%; Score 13; DB 5; Length 24;

Best Local Similarity 100.0%; Pred. No. 7; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPLVC 163

DB 1 DSCQDSDGGPLVC 13

## RESULT 15

AAE17239  
 ID AAE17239 standard; peptide; 35 AA.

XX AAE17239;  
AC  
XX  
XX  
XX 18-APR-2002 (first entry)  
XX Human transmembrane serine protease-related apple domain #1.  
XX  
XX Transmembrane serine protease; gene therapy; metastasis; tumour;  
KW chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;  
KW atherosclerosis; neurodegenerative disease; neuroprotective; cytostatic;  
KW pathogenic infection; antiinflammatory; antiarteriosclerotic;  
KW antibacterial; apple domain.  
XX  
XX Unidentified.  
OS  
XX  
XX WO200196538-A2.  
XX  
XX 20-DEC-2001.  
PD  
XX  
XX 12-JUN-2001; 2001WO-EP006618.  
PF  
XX  
XX 13-JUN-2000; 2000US-0211224P.  
PR  
XX 13-APR-2001; 2001US-0283353P.  
PR  
XX 16-APR-2001; 2001US-0283648P.  
PR  
XX (FARB ) BAYER AG.  
PA  
XX  
XX Xiao Y, Gedrich R;  
PI  
XX  
XX WPI; 2002-098065/13.  
DR  
XX  
XX Novel isolated polynucleotide encoding transmembrane serine protease  
PT polypeptide, for treating chronic obstructive pulmonary disease, tumor  
PT angiogenesis, inflammation, atherosclerosis and neurodegenerative  
PT disease.  
PT  
XX  
XX Disclosure; Fig 3; 120pp; English.  
PS  
XX  
XX The present invention relates to an isolated polynucleotide encoding a  
CC transmembrane serine protease polypeptide. Transmembrane serine protease  
CC gene is useful in gene therapy. The invention also relates to a  
CC gene is useful in gene therapy. The invention also relates to a  
CC pharmaceutical composition which is useful for modulating the activity of  
CC transmembrane serine protease in a disease, such as chronic obstructive  
CC pulmonary disease (COPD), metastasis of malignant cells, tumour  
CC angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease  
CC or pathogenic infection. Transmembrane serine protease is useful as a  
CC bait protein in a two-hybrid or three-hybrid assay. The polypeptide is  
CC useful for generating antibodies against it and in various assay systems.  
CC The present sequence is a human transmembrane serine protease-related  
CC apple domain  
CC  
XX  
XX Sequence 35 AA;  
SQ

Query Match 6.3%; Score 13; DB 5; Length 35;  
Best Local Similarity 100.0%; Pred. No. 8.6;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 DSCQDSGGPLVC 163  
|||||  
Db 9 DSCQDSGGPLVC 21

Search completed: October 27, 2004, 09:29:18  
Job time : 109 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:20:22 ; Search time 28 Seconds  
(without alignments)  
490.279 Million cell updates/sec

Title: US-09-992-095B-54  
Perfect score: 207  
Sequence: 1 MHFCGTLISPEWVLTAAHC.....GVYVRVSRFTWIEGWMRNN 207

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 478139 seqs, 66318000 residues

Word size : 0

Total number of hits satisfying chosen parameters: 374783

Minimum DB seq length: 0  
Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/prodata/1/iaa/6C COMB.pcp.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	207	100.0	207	4	US-10-000-489-54
2	42	20.3	197	1	US-08-456-840-48
3	42	20.3	197	1	US-08-266-407A-48
4	42	20.3	197	2	US-08-892-544-48
5	13	6.3	23	4	US-09-879-792-22
6	13	6.3	149	3	US-09-518-046-20
7	13	6.3	151	3	US-09-518-046-23
8	12	5.8	151	4	US-09-270-767-33178
9	12	5.8	151	3	US-09-270-767-48395
10	12	5.8	159	3	US-09-518-046-24
11	11	5.3	12	4	US-09-879-792-16
12	11	5.3	84	3	US-08-906-769-99
13	11	5.3	84	3	US-08-906-616-99
14	11	5.3	84	3	US-08-817-795-99
15	11	5.3	84	3	US-08-639-075A-99
16	11	5.3	84	3	US-09-012-431-99
17	11	5.3	84	3	US-09-012-692-99
18	11	5.3	84	3	US-08-906-613-99
19	11	5.3	84	5	PCT-US95-14442A-99
20	11	5.3	138	6	5200340-4
21	11	5.3	162	4	US-09-244-111-6
22	11	5.3	207	4	US-09-244-111-4
23	10	4.8	42	1	US-08-293-778-14
24	10	4.8	151	3	US-09-518-046-21
25	10	4.8	154	3	US-09-261-416-5
26	10	4.8	156	3	US-09-261-416-6
27	10	4.8	182	4	US-09-328-925-12

28	10	4.8	200	3	US-09-008-271A-5	Sequence 5, Appli
29	9	4.3	9	1	US-07-819-361-5	Sequence 5, Appli
30	9	4.3	9	1	US-08-179-574-5	Sequence 5, Appli
31	9	4.3	9	3	US-09-518-046-58	Sequence 58, Appl
32	9	4.3	9	4	US-09-618-259-27	Sequence 27, Appl
33	9	4.3	9	5	PCT-US93-00325-5	Sequence 5, Appli
34	9	4.3	20	1	US-08-472-228A-19	Sequence 19, Appl
35	9	4.3	20	3	US-09-146-831-19	Sequence 19, Appl
36	9	4.3	20	5	PCT-US96-09303-19	Sequence 19, Appl
37	9	4.3	22	1	US-08-392-828C-31	Sequence 31, Appl
38	9	4.3	22	3	US-09-330-945-31	Sequence 31, Appl
39	9	4.3	23	4	US-09-879-792-15	Sequence 15, Appl
40	9	4.3	36	3	US-08-944-483-26	Sequence 26, Appl
41	9	4.3	55	4	US-08-843-076D-22	Sequence 22, Appl
42	9	4.3	86	1	US-08-485-455D-53	Sequence 53, Appl
43	9	4.3	86	2	US-08-482-130C-53	Sequence 53, Appl
44	9	4.3	86	2	US-08-484-211C-53	Sequence 53, Appl
45	9	4.3	86	3	US-08-906-769-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1  
US-10-000-489-54  
; Sequence 54, Application US/10000489  
; Patent No. 6794363  
; GENERAL INFORMATION:  
; APPLICANT: Benjanin, Stephane  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US6.DIV  
; CURRENT APPLICATION NUMBER: US 10/000,489  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-54

Query Match	100.0%;	Score 207;	DB 4;	Length 207;
Best Local Similarity	100.0%;	Pred. No. 6.2e-210;		
Matches 207;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	1	MHFCGTLISPEWVLTAAHCLEKSPSSYKVLGHAEVNLPHVQIEVSRLEPTR	60	
Db	1	MHFCGTLISPEWVLTAAHCLEKSPSSYKVLGHAEVNLPHVQIEVSRLEPTR	60	
OY	61	KDIALKLSSPAVITDKVIPACLPSPNYVADRTECFITGWGETQGTGAGLLKEAQLPV	120	
Db	61	KDIALKLSSPAVITDKVIPACLPSPNYVADRTECFITGWGETQGTGAGLLKEAQLPV	120	
OY	121	IENKVCNRYEFLNGRVQSTELCAGHLAAGTDSQQSGGSLVCFEKDKYILQGVTSWGLG	180	
Db	121	IENKVCNRYEFLNGRVQSTELCAGHLAAGTDSQQSGGSLVCFEKDKYILQGVTSWGLG	180	
OY	181	CARPNGKGVYVRVSRFTWIEGWMRNN	207	
Db	181	CARPNGKGVYVRVSRFTWIEGWMRNN	207	

RESULT 2  
US-08-456-840-48  
; Sequence 48, Application US/08456840  
; Patent No. 5597908  
; GENERAL INFORMATION:  
; APPLICANT: Taddei-Peters, W. C.  
; APPLICANT: Butler, Sandra M.  
; TITLE OF INVENTION: Immunoactive Peptides of Apo(a)  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5597908el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,840  
; FILING DATE: 01-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/266,407  
; FILING DATE: 27-JUN-1994  
; APPLICATION NUMBER: US 08/172,461  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gormley, Mary E.  
; REGISTRATION NUMBER: 34,409  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-258-5200  
; TELEFAX: 301-977-0847  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 197 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; US-08-456-840-48

Query Match 20.3%; Score 42; DB 1; Length 197;  
Best Local Similarity 100.0%; Pred. No. 4e-36;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 GTDSCQSGGGLVCFEKKYILQGVTSWGLGCARPKNKPGVY 190  
|||||  
Db 141 GTDSCQSGGGLVCFEKKYILQGVTSWGLGCARPKNKPGVY 182  
|||||

RESULT 3  
US-08-266-407A-48  
; Sequence 48, Application US/08266407A  
; Patent No. 5786156  
; GENERAL INFORMATION:  
; APPLICANT: Taddei-Peters, W. C.  
; APPLICANT: Butler, Sandra M.  
; TITLE OF INVENTION: Immunoactive Peptides of Apo(a)  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5786156el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/892,544  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/266,407  
; FILING DATE: 27-JUN-1994  
; APPLICATION NUMBER: US 08/172,461  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gormley, Mary E.  
; REGISTRATION NUMBER: 34,409  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-258-5200  
; TELEFAX: 301-977-0847  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 197 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; US-08-456-840-48

Query Match 20.3%; Score 42; DB 1; Length 197;  
Best Local Similarity 100.0%; Pred. No. 4e-36;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 GTDSCQSGGGLVCFEKKYILQGVTSWGLGCARPKNKPGVY 190  
|||||  
Db 141 GTDSCQSGGGLVCFEKKYILQGVTSWGLGCARPKNKPGVY 182  
|||||

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/266,407A  
; FILING DATE: 27-JUN-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/172,461  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gormley, Mary E.  
; REGISTRATION NUMBER: 34,409  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-258-5200  
; TELEFAX: 301-977-0847  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 197 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; US-08-266-407A-48

Query Match 20.3%; Score 42; DB 1; Length 197;  
Best Local Similarity 100.0%; Pred. No. 4e-36;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 149 GTDSCQSGGGLVCFEKKYILQGVTSWGLGCARPKNKPGVY 190  
|||||  
Db 141 GTDSCQSGGGLVCFEKKYILQGVTSWGLGCARPKNKPGVY 182  
|||||

RESULT 4  
US-08-892-544-48  
; Sequence 48, Application US/08892544  
; Patent No. 5874544  
; GENERAL INFORMATION:  
; APPLICANT: Taddei-Peters, W. C.  
; APPLICANT: Butler, Sandra M.  
; TITLE OF INVENTION: Immunoactive Peptides of Apo(a)  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5874544el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/892,544  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/266,407  
; FILING DATE: 27-JUN-1994  
; APPLICATION NUMBER: US 08/172,461  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gormley, Mary E.  
; REGISTRATION NUMBER: 34,409  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-258-5200  
; TELEFAX: 301-977-0847

INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-892-544-48

Query Match 20.3%; Score 42; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 4e-36;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 149 GTDSCQDGGPLVCFEKDYKYLQGVTSWGLGCARPKNKGVY 190  
DB 141 GTDSCQDGGPLVCFEKDYKYLQGVTSWGLGCARPKNKGVY 182

RESULT 5  
US-09-879-792-22  
Sequence 22, Application US/09879792  
Patent No. 6734006  
GENERAL INFORMATION:  
APPLICANT: Xiao, Yonghong  
APPLICANT: Gedrich, Richard  
TITLE OF INVENTION: Regulation of Human Transmembrane Serine  
FILE REFERENCE: 02973.00035  
CURRENT APPLICATION NUMBER: US/09/879,792  
CURRENT FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/211,224  
PRIOR FILING DATE: 2000-06-13  
PRIOR APPLICATION NUMBER: US 60/283,353  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: US 60/283,648  
PRIOR FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: PCT  
PRIOR FILING DATE: 2001-06-12  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: BLOCKS BL00134B  
US-09-879-792-22 (Docket No. 6734006 LIO-81-WO)

Query Match 6.3%; Score 13; DB 4; Length 23;  
Best Local Similarity 100.0%; Pred. No. 1.9e-06;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 DSCQDGGPLVC 163  
DB 1 DSCQDGGPLVC 13

RESULT 6  
US-09-518-046-20  
Sequence 20, Application US/09518046  
Patent No. 6294663  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
APPLICANT: Underwood, Lowell J.  
TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
FILE REFERENCE: D6192CIP  
CURRENT APPLICATION NUMBER: US/09/518,046  
CURRENT FILING DATE: 2000-03-02  
EARLIER APPLICATION NUMBER: 09/261,416  
EARLIER FILING DATE: 1999-03-03  
NUMBER OF SEQ ID NOS: 153

SEQ ID NO 20  
LENGTH: 149  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
OTHER INFORMATION: protease domain of protease M (Prom)  
US-09-518-046-20

Query Match 6.3%; Score 13; DB 3; Length 149;  
Best Local Similarity 100.0%; Pred. No. 1.1e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 DSCQDGGPLVC 163  
DB 137 DSCQDGGPLVC 149

RESULT 7  
US-09-518-046-23  
Sequence 23, Application US/09518046  
Patent No. 6294663  
GENERAL INFORMATION:  
APPLICANT: O'Brien, Timothy J.  
APPLICANT: Underwood, Lowell J.  
TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
FILE REFERENCE: D6192CIP  
CURRENT APPLICATION NUMBER: US/09/518,046  
CURRENT FILING DATE: 2000-03-02  
EARLIER APPLICATION NUMBER: 09/261,416  
EARLIER FILING DATE: 1999-03-03  
NUMBER OF SEQ ID NOS: 153  
SEQ ID NO 23  
LENGTH: 157  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
OTHER INFORMATION: protease domain of TADG-12 (TADG12)  
US-09-518-046-23

Query Match 6.3%; Score 13; DB 3; Length 157;  
Best Local Similarity 100.0%; Pred. No. 1.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 151 DSCQDGGPLVC 163  
DB 145 DSCQDGGPLVC 157

RESULT 8  
US-09-270-767-33178  
Sequence 33178, Application US/09270767  
Patent No. 6703491  
GENERAL INFORMATION:  
APPLICANT: Homburger et al.  
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
FILE REFERENCE: File Reference: 7326-094  
CURRENT APPLICATION NUMBER: US/09/270,767  
CURRENT FILING DATE: 1999-03-17  
NUMBER OF SEQ ID NOS: 62517  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 33178  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Drosophila melanogaster  
FEATURE:  
OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-33178  
Query Match 5.8%; Score 12; DB 4; Length 151;  
Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 DSCQDGGGGLV 162  
|  
Db 72 DSCQDGGGGLV 83

## RESULT 9

US-09-270-767-48395  
; Sequence 48395, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48395  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-48395

Query Match 5.8%; Score 12; DB 4; Length 151;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 DSCQDGGGGLV 162  
|  
Db 72 DSCQDGGGGLV 83

## RESULT 10

US-09-518-046-24  
; Sequence 24, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 24  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: protease domain of TMPRSS2 (Tmprss2)  
US-09-518-046-24

Query Match 5.8%; Score 12; DB 3; Length 159;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 DSCQDGGGGLV 162  
|  
Db 147 DSCQDGGGGLV 158

## RESULT 11

US-09-879-792-16  
; Sequence 16, Application US/09879792  
; Patent No. 6734006

; GENERAL INFORMATION:  
; APPLICANT: Xiao, Yonghong  
; APPLICANT: Gedrich, Richard  
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine  
; FILE REFERENCE: 02973.00035  
; CURRENT APPLICATION NUMBER: US/09/879,792  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/211,224  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US 60/283,353  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/283,648  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: BLOCKS BL1253G  
US-09-879-792-16

Query Match 5.3%; Score 11; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 153 CQGDGGGGLV 163  
|  
Db 2 CQGDGGGGLV 12

## RESULT 12

US-08-906-769-99  
; Sequence 99, Application US/08906769  
; Patent No. 6077687  
; GENERAL INFORMATION:  
; APPLICANT: Grieve, Robert B.  
; APPLICANT: Rushlow, Keith E.  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Frank, Glenn R.  
; APPLICANT: Stiegler, Gary  
; APPLICANT: Gaines, Patrick J.  
; APPLICANT: Silver, Gary  
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/08/906,769  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 24  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: protease domain of TMPRSS2 (Tmprss2)  
US-08-906-769-99

US-09-879-792-16  
; Sequence 16, Application US/09879792  
; Patent No. 6734006



TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223 99:  
INFORMATION FOR SEQ ID NO: 99:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-906-769-99

Query Match 5.3%; Score 11; DB 3; Length 84;  
Best Local Similarity 100.0%; Pred. No. 0.00083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPL 161  
Db 74 DSCQDSDGGPL 84

RESULT 13  
US-08-906-616-99

Sequence 99, Application US/08906616  
Patent No. 6121035

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.  
APPLICANT: Rushlow, Keith E.  
APPLICANT: Wu Hunter, Shirley  
APPLICANT: Frank, Glenn R.  
APPLICANT: Stiegler, Gary  
APPLICANT: Gaines, Patrick J.  
APPLICANT: Silver, Gary

TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF

NUMBER OF SEQUENCES: 190  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/906,616  
FILING DATE: 05-AUG-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-25-C2-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 99:

SEQUENCE CHARACTERISTICS:  
LENGTH: 84 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-906-616-99

Query Match 5.3%; Score 11; DB 3; Length 84;  
Best Local Similarity 100.0%; Pred. No. 0.00083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPL 161  
Db 74 DSCQDSDGGPL 84

RESULT 14

US-08-817-795-99

Sequence 99, Application US/08817795  
Patent No. 6139840

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.  
APPLICANT: Rushlow, Keith E.  
APPLICANT: Hunter, Shirley Wu  
APPLICANT: Frank, Glenn R.  
APPLICANT: Heath, Andrew W.  
APPLICANT: Yamaka, Miles Yamanaka  
APPLICANT: Arfsten, Ann  
APPLICANT: Dale, Beverly  
APPLICANT: Stiegler, Gary

TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND

TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA  
INFECTION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID  
MOLECULES, AND USES THEREOF

NUMBER OF SEQUENCES: 119

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/817,795

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/14442

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Gary J. Connell

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 99:

SEQUENCE CHARACTERISTICS:

LENGTH: 84 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-817-795-99

Query Match 5.3%; Score 11; DB 3; Length 84;  
Best Local Similarity 100.0%; Pred. No. 0.00083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPL 161  
Db 74 DSCQDSDGGPL 84

RESULT 15

US-08-639-075A-99

Sequence 99, Application US/08639075A  
Patent No. 6150125

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.

APPLICANT: Rushlow, Keith E.

APPLICANT: Wu Hunter, Shirley

APPLICANT: Frank, Glenn R.

APPLICANT: Stiegler, Gary

Thu Oct 28 07:18:43 2004

;; APPLICANT: Gaines, Patrick J.  
;; APPLICANT: Silver, Gary  
;; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
;; TITLE OF INVENTION: MOLECULES AND USES THEREOF  
;; NUMBER OF SEQUENCES: 190  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sheridan Ross & McIntosh  
;; STREET: 1700 Lincoln Street, Suite 3500  
;; CITY: Denver  
;; STATE: Colorado  
;; COUNTRY: USA  
;; ZIP: 80203  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/639,075A  
;; FILING DATE: 24-APR-1996  
;; CLASSIFICATION: 424  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Connell, Gary J.  
;; REGISTRATION NUMBER: 32,020  
;; REFERENCE/DOCKET NUMBER: 2618-25-C2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (303) 863-9700  
;; TELEFAX: (303) 863-0223  
;; INFORMATION FOR SEQ ID NO: 99:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 84 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-639-075A-99

Query Match 5.3%; Score 11; DB 3; Length 84;  
Best Local Similarity 100.0%; Pred.No. 0.00083;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 DSCQDSDSGGPL 161  
|||||  
Db 74 DSCQDSDSGGPL 84

Search completed: October 27, 2004, 09:33:09  
Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 27, 2004, 09:33:13 ; Search time 4076 Seconds  
(without alignments)  
2401.611 Million cell updates/sec

Title: US-09-992-095B-54  
Perfect score: 207  
Sequence: 1 MHFCGGLTSPWVLTAAHC.....GVYVRVSRFTWIEGWRNN 207

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues  
Word size: 1

Total number of hits satisfying chosen parameters: 5947187

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DBV=xlh  
-Q/cgn2 1/USPTO.spool/US0992095/runat\_26102004\_090252\_11801/app.query.fasta\_1.391  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=621  
-USER=US0992095 @CGN\_1\_1\_3731 @runat\_26102004\_090252\_11801 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	25.6	356	6	CQ433162 Sequence
2	53	25.6	436	6	CQ506843 Sequence
3	53	25.6	440	6	CQ476920 Sequence
4	53	25.6	470	6	CQ425554 Sequence

5	49	23.7	254	9	HUMPLG23	M34275 Human plasm
6	49	23.7	468	6	CQ415403	CQ415403 Sequence
7	47	22.7	567	4	BOVPMG	K02935 Bovine plas
8	46	22.2	331	9	HUMPLG20	M34272 Human plasm
9	43	20.8	586	11	BV100085	BV100086 RPAMWSEQ0
10	43	20.8	592	11	BV159185	BV159185 RPAMWSEQ0
11	37	17.9	251	9	HUMPLG21	M34273 Human plasm
12	27	13.0	312	6	AX418027	AX418027 Sequence
13	27	13.0	312	6	AX978721	AX978721 Sequence
14	27	13.0	312	6	BD113580	BD113580 EST and e
15	23	11.1	361	9	HUMPLG19	M3290 Human plasm
16	20	9.7	60	6	AX611981	AX611981 Sequence
17	20	9.7	60	6	AX611983	AX611983 Sequence
18	17	8.2	252	6	CQ439589	CQ439589 Sequence
19	16	7.7	48	6	AX611977	AX611977 Sequence
20	16	7.7	48	6	AX611979	AX611979 Sequence
21	16	7.7	591	11	BV100085	BV100085 RPAMWSEQ0
22	16	7.7	591	11	BV159184	BV159184 RPAMWSEQ0
23	14	6.8	44	6	AX611985	AX611985 Sequence
24	14	6.8	375	9	HXS139	Z36790 H.sapiens (
25	14	6.8	430	6	AX778535	AX778535 Sequence
26	14	6.8	529	4	AY455801	AY455801 Canis fam
27	13	6.3	181	6	CQ408336	CQ408336 Sequence
28	13	6.3	201	11	BV207480	BV207480 sqmm22359
29	13	6.3	201	11	BV207482	BV207482 sqmm22360
30	13	6.3	262	6	AR152171	AR152171 Sequence
31	13	6.3	296	6	CQ731830	CQ731830 Sequence
32	13	6.3	425	6	CQ395619	CQ395619 Sequence
33	13	6.3	425	6	CQ401952	CQ401952 Sequence
34	13	6.3	439	6	AX277349	AX277349 Sequence
35	13	6.3	439	6	AX277375	AX277375 Sequence
36	13	6.3	479	6	CQ395146	CQ395146 Sequence
37	13	6.3	479	6	CQ401486	CQ401486 Sequence
38	13	6.3	498	6	CQ407870	CQ407870 Sequence
39	13	6.3	546	6	AX775014	AX775014 Sequence
40	13	6.3	583	11	BV159179	BV159179 RPAMWSEQ0
41	13	6.3	614	6	AX369345	AX369345 Sequence
42	12	5.8	213	6	A27460	A27460 oligonucleo
43	12	5.8	260	11	G38616	G38616 SHGC-63628
44	12	5.8	304	6	AX524081	AX524081 Sequence
45	12	5.8	304	6	AX552819	AX552819 Sequence

ALIGNMENTS

RESULT 1	CQ433162	356 bp	DNA	linear	PAT 28-JAN-2004
LOCUS	Sequence 18196 from Patent WO0151628.				
DEFINITION	CQ433162				
ACCESSION	CQ433162.1				
VERSION	CQ433162.1				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.				
TITLE	Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer				
JOURNAL	Patent: WO 0151628-A 18196 19-JUL-2001;				
FEATURES	Millennium Pharmaceuticals, Inc. (US)				
source	Location/Qualifiers				
	1..356				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN	Alignment Scores:	9.4e-45	Length:	356
	Pred. No.:	53.00	Matches:	53
	Score:	100.00%	Conservative:	0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.60% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CQ433162 (1-356)

Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174
Db 152 GGTGACAGTGGAGGCCCTCTGGTTGCTTTGAGAGGACAAATACATTTTCAAGGAGTC 211
Qy 175 ThrSerTrpGlyLeuGlyCysAlaArgProAenLysProGlyValTyrValArgValSer 194
Db 212 ACTTCTGGGGTCTGGCTGTGCAGCCCCCAATAGCCTGGTGTCTATGTTGCTGTTTCA 271
Qy 195 ArgPheValThrTrpIleGluGlyValMetArgAenAen 207
Db 272 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 310

RESULT 2
CQ506843 Sequence 38710 from Patent WO0160860. linear PAT 30-JAN-2004
LOCUS CQ506843
DEFINITION CQ506843
ACCESSION CQ506843
VERSION CQ506843.1 GI:41473029
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 38710 23-AUG-2001;
Mileennium Predictive Medicine, Inc. (US)
FEATURES
Location/Qualifiers
source 1..436
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.11e-44 Length: 436 bp DNA linear PAT 30-JAN-2004
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.60% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CQ506843 (1-436)

Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174
Db 199 GGTGACAGTGGAGGCCCTCTGGTTGCTTTGAGAGGACAAATACATTTTCAAGGAGTC 258
Qy 175 ThrSerTrpGlyLeuGlyCysAlaArgProAenLysProGlyValTyrValArgValSer 194
Db 259 ACTTCTGGGGTCTGGCTGTGCAGCCCCCAATAGCCTGGTGTCTATGTTGCTGTTTCA 318
Qy 195 ArgPheValThrTrpIleGluGlyValMetArgAenAen 207
Db 319 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 357

RESULT 3
CQ476920 Sequence 8787 from Patent WO0160860. linear PAT 30-JAN-2004
LOCUS CQ476920
DEFINITION CQ476920
ACCESSION CQ476920
VERSION CQ476920.1 GI:41442539
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 10588 19-JUL-2001;
Mileennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source 1..470
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.18e-44 Length: 470
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.60% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CQ425554 (1-470)

Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174

```

```

REFERENCE
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 8787 23-AUG-2001;
Mileennium Predictive Medicine, Inc. (US)
FEATURES
Location/Qualifiers
source 1..440
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.12e-44 Length: 440
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.60% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CQ476920 (1-440)

Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174
Db 198 GGTGACAGTGGAGGCCCTCTGGTTGCTTTGAGAGGACAAATACATTTTCAAGGAGTC 257
Qy 175 ThrSerTrpGlyLeuGlyCysAlaArgProAenLysProGlyValTyrValArgValSer 194
Db 258 ACTTCTGGGGTCTGGCTGTGCAGCCCCCAATAGCCTGGTGTCTATGTTGCTGTTTCA 317
Qy 195 ArgPheValThrTrpIleGluGlyValMetArgAenAen 207
Db 318 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 356

RESULT 4
CQ425554 Sequence 10588 from Patent WO0151628. linear PAT 28-JAN-2004
LOCUS CQ425554
DEFINITION CQ425554
ACCESSION CQ425554
VERSION CQ425554.1 GI:41377783
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 10588 19-JUL-2001;
Mileennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source 1..470
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.18e-44 Length: 470
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.60% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CQ425554 (1-470)

Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174

```

Db 156 GGTGACAGTGGAGGCGCTCTGGTTTGTGAGAGGACAAATACATATTTTACAAGGAGTC 215

Qy 175 ThrSerTrpGlyLeuGlyCysAlaArgProAenLysProGlyValTrValArgValSer 194  
|||||  
Db 216 ACTTCTGGGGCTTGGCTGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTGTTTCA 275  
|||||

Qy 195 ArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207  
|||||  
Db 276 AGGTTGTACTTGGATTGAGGAGTGATGAGAAATAAT 314  
|||||

RESULT 5  
LOCUS HUMPLG23 254 bp DNA linear PRI 07-JAN-1995  
DEFINITION Human plasminogen gene, exon 18.  
ACCESSION M34275 J05286  
VERSION M34275.1 GI:190063  
KEYWORDS plasminogen.  
SEGMENT 23 of 24  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 254)  
AUTHORS Petersen,T.E., Martzen,M.R., Ichinose,A. and Davie,E.W.  
TITLE Characterization of the gene for human plasminogen, a key proenzyme  
in the fibrinolytic system  
J. Biol. Chem. 265 (11), 6104-6111 (1990)  
JOURNAL 90202879  
MEDLINE 2318848  
PUBMED  
COMMENT Original source text: Homo sapiens DNA.  
Draft entry and computer-readable sequence for [1] kindly submitted  
by A. Ichinose, 26-MAR-1990, for release after publication.

FEATURES  
source  
1..254  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="6q26-q27"  
/cell\_type="leukocyte and lung fibroblast"  
intron <1..68  
/genes="PLG"  
/note="G00-119-498"  
/number=17  
69..214  
/genes="PLG"  
/note="G00-119-498"  
/number=18  
215..>254  
/genes="PLG"  
/note="G00-119-498"  
/number=18  
intron  
Undetermined number of base pairs after segment 22.

ALIGNMENT Scores:  
Pred. No.: 9,07e-41 Length: 254  
Score: 49.00 Matches: 49  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 23.67% Indels: 0  
DB: 9 Gaps: 0

US-09-992-095B-54 (1-207) x HUMPLG23 (1-254)

Qy 106 GlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProValIleGluAenLysVal 125  
|||||  
Db 68 GGTACTTTGGAGCTGGCTTCTCAAGGAAGCCAGCTCCCTGTGATTGAGATAAAGTG 127  
|||||

Qy 126 CysAsnArgTyrGluPheLeuAenGlyArgValGlnSerThrGluLeuCysAlaGlyHis 145  
|||||  
Db 128 TGAATCGCTATGAGTTCTGAAATGGAAGAGTCCAATCCACCACTCTGTGTGGGCAT 187  
|||||

Qy 146 LeuAlaGlyGlyThrAspSerCysGln 154  
|||||

Db 188 TTGGCCCGAGGCACTGACAGTTGGCCAG 214

RESULT 6  
LOCUS CQ415403 468 bp DNA linear PAT 28-JAN-2004  
DEFINITION Sequence 437 from Patent WO0151628.  
ACCESSION CQ415403  
VERSION CQ415403.1 GI:41367632  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.  
TITLE Novel genes, compositions, kits, and methods for identification,  
assessment, prevention, and therapy of breast cancer  
JOURNAL Patent: WO 0151628-A 437 19-JUL-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..468  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.5e-40 Length: 468  
Score: 49.00 Matches: 49  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 23.67% Indels: 0  
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CQ415403 (1-468)

Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174  
|||||  
Db 171 GGTGACAGTGGAGGCGCTCTGGTTTGTGAGAGGACAAATACATATTTTACAAGGAGTC 230  
|||||

Qy 175 ThrSerTrpGlyLeuGlyCysAlaArgProAenLysProGlyValTrValArgValSer 194  
|||||  
Db 231 ACTTCTTGGGGTCTTGGCTGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTGTTTCA 290  
|||||

Qy 195 ArgPheValThrTrpIleGluGlyVal 203  
|||||  
Db 291 AGTTTGTACTTGGATTGAGGAGTT 317  
|||||

RESULT 7  
LOCUS BOVPMG 567 bp mRNA linear MAM 27-APR-1993  
DEFINITION Bovine plasminogen mRNA, 3' end.  
ACCESSION K02935  
VERSION K02935.1 GI:163551  
KEYWORDS glycoprotein; plasminogen.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovinae; Bos.  
REFERENCE 1 (bases 1 to 567)  
AUTHORS Malinowski,D.P., Sadler,J.E. and Davie,E.W.  
TITLE Characterization of a complementary deoxyribonucleic acid coding  
for human and bovine plasminogen  
JOURNAL Biochemistry 23 (18), 4243-4250 (1984)  
MEDLINE 85023311  
PUBMED 6148961  
COMMENT Original source text: Bovine liver, cDNA to mRNA, clone 5.  
FEATURES  
Location/Qualifiers  
1..567  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"



```

DB:          11          Gaps:          0
US-09-992-095B-54 (1-207) x BV100086 (1-586)

QY 155 GlyAspSerGlyClyProLeuValCysPheGluLysAspIlysTyrIleLeuGlnGlyVal 174
DB 360 GGCACAGTGGAGGACCTCTGGTTGCTTCGAGAAGGACAAGTACATTTTACAAGGAGTC 301

QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyrValArgValSer 194
DB 300 ACTTCTTGGGGTCTTGGCTGTGCTCGCCCCAATAAGCCTGGTGTCTACGTTCTGTCTCA 241

QY 195 ArgPheVal 197
DB 240 CGGTTTGT 232

RESULT 10
BV159185          592 bp      DNA      linear      STS 14-MAY-2004
LOCUS      RPAMSE0032117 Roche Palo Alto Mus musculus STS genomic, sequence
DEFINITION      tagged site.
ACCESSION      BV159185
VERSION        BV159185.1 GI:47174420
KEYWORDS
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 592)
AUTHORS      Usuka, J., Liao, G., Cheng, J., Nguyen, A., Bach, C., Puech, A.,
McPherson, J.D., Foerzler, D. and Peltz, G.
TITLE        Mus musculus SNPs
JOURNAL
COMMENT      Unpublished (2003)

Contact: Jonathan Usuka
Roche Palo Alto Genetics and Genomics Department
Roche Palo Alto
3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
Tel: 6508555807
Email: Jonathan.Usuka@roche.com
Primer A: No primer submitted with this STS
Primer B: No primer submitted with this STS.

FEATURES             source
     1..592
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10090"
     /map="17-8835-8244-AC087901.20.1.214372"
     /clone_lib="Roche Palo Alto"
     /note="SNPs developed from assay sequences derived from 15
different strains of mice (as of October 1, 2003). Those
strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/Osnu,
BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
MRL/MpJ, NZB/BinJ, NZW/LacJ, SPRET/Ei."
     <1..>592

STSS
ORIGIN
Alignment Scores:
Pred. No.:      2,6e-34      Length:      592
Score:          43.00      Matches:      43
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    20.77%      Indels:      0
DB:            11      Gaps:      0

US-09-992-095B-54 (1-207) x BV159185 (1-592)

QY 155 GlyAspSerGlyClyProLeuValCysPheGluLysAspIlysTyrIleLeuGlnGlyVal 174
DB 217 GGCACAGTGGAGGACCTCTGGTTGCTTCGAGAAGGACAAGTACATTTTACAAGGAGTC 276

QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyrValArgValSer 194
DB 300 ACTTCTTGGGGTCTTGGCTGTGCTCGCCCCAATAAGCCTGGTGTCTACGTTCTGTCTCA 241

DB:          11          Gaps:          0
US-09-992-095B-54 (1-207) x HUMPLG21 (1-251)

QY 70 SerProAlaValIleThrAspLysValIleProAlaCysLeuProSerProAsnTyrVal 89
DB 39 AGTCCTGCCGTCATCACTGACAAAGTAATCCAGCTTGTCTGCATCCCCAAATATGTG 98

QY 90 ValAlaAspArgThrGluCysPheIleThrGlyTyrGlyGluThrGlnGly 106
DB 99 GTCGCTGACCGGACCAATGTTTCATCACTGGCTGGGAGAAACCCCAAGGT 149

RESULT 12
AR418027          312 bp      DNA      linear      PAT 18-DEC-2003
LOCUS      AR418027
DEFINITION      Sequence 9524 from patent US 6639063.
ACCESSION      AR418027
VERSION        AR418027.1 GI:40173137

```

```

DB 277 ACTTCTTGGGGTCTTGGCTGTGCTCGCCCCAATAAGCCTGGTGTCTACGTTCTGTCTCA 336

QY 195 ArgPheVal 197
DB 337 CGGTTTGT 345

RESULT 11
HUMPLG21          251 bp      DNA      linear      PRI 07-JAN-1995
LOCUS      Human plasminogen gene, exon 17.
DEFINITION      M34273 J05286
ACCESSION      M34273.1 GI:190061
VERSION        M34273.1
KEYWORDS      plasminogen.
SEGMENT       21 of 24
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 251)
AUTHORS      Petersen, T.E., Martzen, M.R., Ichinose, A. and Davie, E.W.
TITLE        Characterization of the gene for human plasminogen, a key proenzyme
in the fibrinolytic system
JOURNAL      J. Biol. Chem. 265 (11), 6104-6111 (1990)
MEDLINE      90202879
PUBMED       2318848
COMMENT      Original source text: Homo sapiens DNA.
Draft entry and computer-readable sequence for [1] kindly submitted
by A. Ichinose, 26-MAR-1990, for release after publication.

FEATURES             source
     1..251
     /organism="Homo sapiens"
     /mol_type="genomic DNA"
     /db_xref="taxon:9606"
     /map="6q26-q27"
     /cell_type="leukocyte and lung fibroblast"
     <1..40
     /gene="PLG"
     /note="G00-119-498"
     /number=16
     41..147
     /gene="PLG"
     /note="G00-119-498"
     /number=17
     148..>251
     /gene="PLG"
     /note="G00-119-498"
     /number=17

ORIGIN      Undetermined number of base pairs after segment 20.
Alignment Scores:
Pred. No.:      1,85e-28      Length:      251
Score:          37.00      Matches:      37
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    17.87%      Indels:      0
DB:            9      Gaps:      0

US-09-992-095B-54 (1-207) x HUMPLG21 (1-251)

QY 70 SerProAlaValIleThrAspLysValIleProAlaCysLeuProSerProAsnTyrVal 89
DB 39 AGTCCTGCCGTCATCACTGACAAAGTAATCCAGCTTGTCTGCATCCCCAAATATGTG 98

QY 90 ValAlaAspArgThrGluCysPheIleThrGlyTyrGlyGluThrGlnGly 106
DB 99 GTCGCTGACCGGACCAATGTTTCATCACTGGCTGGGAGAAACCCCAAGGT 149

RESULT 12
AR418027          312 bp      DNA      linear      PAT 18-DEC-2003
LOCUS      AR418027
DEFINITION      Sequence 9524 from patent US 6639063.
ACCESSION      AR418027
VERSION        AR418027.1 GI:40173137

```

```

KEYWORDS          Unknown.
SOURCE             Unknown.
ORGANISM           Unclassified.
REFERENCE          1 (bases 1 to 312)
AUTHORS            Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE             EST's and encoded human proteins
JOURNAL           Patent: US 6639063-A 9524 28-OCT-2003;
FEATURES           Location/Qualifiers
                    source
                    1..312
                    /organism="unknown"
                    /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
  Pred. No.:      4.03e-18      Length:      312
  Score:          27.00         Matches:     27
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:    13.04%        Indels:      0
  DB:             6            Gaps:        0

US-09-992-095B-54 (1-207) x AR418027 (1-312)

QY 122 GluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeu 141
    |||||
Db 2 GAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAAATCCACCGAATC 61

QY 142 CysAlaGlyHisLeuAlaGly 148
    |||||
Db 62 TGTGCTGGGCATTGGCCGGA 82

RESULT 13
AX978721
LOCUS             AX978721             312 bp      DNA      linear      PAT 15-JAN-2004
DEFINITION       Sequence 9524 from Patent EP1104808.
ACCESSION        AX978721
VERSION          AX978721.1 GI:40984861
KEYWORDS         Homo sapiens (human)
ORGANISM         Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1
AUTHORS          Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.
TITLE            ESTs and encoded human proteins
JOURNAL          Patent: EP 1104808-A 9524 06-JUN-2001;
FEATURES         Location/Qualifiers
                  source
                  1..312
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
  Pred. No.:      4.03e-18      Length:      312
  Score:          27.00         Matches:     27
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:    13.04%        Indels:      0
  DB:             6            Gaps:        0

US-09-992-095B-54 (1-207) x AX978721 (1-312)

QY 122 GluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeu 141
    |||||
Db 2 GAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAAATCCACCGAATC 61

QY 142 CysAlaGlyHisLeuAlaGly 148
    |||||
Db 62 TGTGCTGGGCATTGGCCGGA 82

KEYWORDS          Unknown.
SOURCE             Unknown.
ORGANISM           Unclassified.
REFERENCE          1 (bases 1 to 312)
AUTHORS            Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.
TITLE             EST's and encoded human proteins
JOURNAL           Patent: US 6639063-A 9524 28-OCT-2003;
FEATURES           Location/Qualifiers
                    source
                    1..312
                    /organism="unknown"
                    /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
  Pred. No.:      4.03e-18      Length:      312
  Score:          27.00         Matches:     27
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:    13.04%        Indels:      0
  DB:             6            Gaps:        0

US-09-992-095B-54 (1-207) x AR418027 (1-312)

QY 122 GluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeu 141
    |||||
Db 2 GAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAAATCCACCGAATC 61

QY 142 CysAlaGlyHisLeuAlaGly 148
    |||||
Db 62 TGTGCTGGGCATTGGCCGGA 82

RESULT 14
BD113580
LOCUS             BD113580             312 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION       EST and encoded human protein.
ACCESSION        BD113580
VERSION          BD113580.1 GI:23208398
KEYWORDS         Homo sapiens (human)
ORGANISM         Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1 (bases 1 to 312)
AUTHORS          Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE            EST and encoded human protein
JOURNAL          Patent: JP 2002010789-A 5657 15-JAN-2002;
FEATURES         Location/Qualifiers
                    source
                    1..312
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
  Pred. No.:      4.03e-18      Length:      312
  Score:          27.00         Matches:     27
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:    13.04%        Indels:      0
  DB:             6            Gaps:        0

US-09-992-095B-54 (1-207) x BD113580 (1-312)

QY 122 GluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeu 141
    |||||
Db 2 GAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAAATCCACCGAATC 61

QY 142 CysAlaGlyHisLeuAlaGly 148
    |||||
Db 62 TGTGCTGGGCATTGGCCGGA 82

RESULT 15
HUMPLG19
LOCUS             HUMPLG19             361 bp      DNA      linear      PRI 07-JAN-1995
DEFINITION       Human plasminogen gene, exon 15.
ACCESSION        M33290 J05286
VERSION          M33290.1 GI:190059
KEYWORDS         plasminogen.
JOURNAL          Segment 19 of 24
FEATURES         Location/Qualifiers
                    source
                    1..312
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
  Pred. No.:      4.03e-18      Length:      312
  Score:          27.00         Matches:     27
  Percent Similarity: 100.00%    Conservative: 0
  Best Local Similarity: 100.00% Mismatches:    0
  Query Match:    13.04%        Indels:      0
  DB:             6            Gaps:        0

US-09-992-095B-54 (1-207) x AX978721 (1-312)

QY 122 GluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeu 141
    |||||
Db 2 GAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAAATCCACCGAATC 61

QY 142 CysAlaGlyHisLeuAlaGly 148
    |||||
Db 62 TGTGCTGGGCATTGGCCGGA 82

REFERENCE        1
AUTHORS          Petersen,T.E., Martzen,M.R., Ichinose,A. and Davie,E.W.
TITLE            Characterization of the gene for human plasminogen, a key proenzyme
                  in the fibrinolytic system
JOURNAL          J. Biol. Chem. 265 (11), 6104-6111 (1990)
MEDLINE          90202879

```



PUBMED 2318848  
COMMENT Original source text: Homo sapiens DNA.  
Draft entry and computer-readable sequence for [1] kindly submitted  
by A. Ichinose, 26-MAR-1990, for release after publication.

FEATURES  
Source  
1. .361  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="eq26-q27"  
/cell\_type="leukocyte and lung fibroblast"  
intron  
1. .65  
/gene="PLG"  
/note="G00-119-498"  
/number=14  
66. .140  
/gene="PLG"  
/note="G00-119-498"  
/number=15  
141. .>361  
intron  
/gene="PLG"  
/note="G00-119-498"  
/number=15  
ORIGIN Undetermined number of base pairs after segment 18.

Alignment Scores:  
Pred. No.: 5.77e-14 Length: 361  
Score: 23.00 Matches: 23  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.11% Indels: 0  
DB: 9 Gaps: 0

US-09-992-095B-54 (1-207) x HUMPLG19 (1-361)

Qy 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20  
Db 73 ATGCACCTTCGTGGAGGCACCTTGATATCCACAGAGTGGGTGTGACTGCTGCCACTGC 132  
Qy 21 LeuGlulys 23  
Db 133 TTGGAGAAG 141

Search completed: October 27, 2004, 10:56:49  
Job time : 4082 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 27, 2004, 09:32:38 ; Search time 503 Seconds

(without alignments)  
2160.299 Million cell updates/sec

Title: US-09-992-095b-54

Perfect score: 207

Sequence: 1 MHFCGGLTSPWVLTAAHC.....GVVVRVSRPVTWIEGVMRNN 207

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6419606

Minimum DB seq length: 0

Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US0992095/runat\_26102004\_090252\_11793/app\_query.fasta\_1.391  
-DB=N Geneseq 23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcio -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=621  
-USER=US0992095.cgn\_1\_1\_470 @runat\_26102004\_090252\_11793 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_23Sep04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2002bs: \*  
8: Geneseqn2003as: \*  
9: Geneseqn2003bs: \*  
10: Geneseqn2003cs: \*  
11: Geneseqn2003ds: \*  
12: Geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	25.6	356	4 AAL25726	Aal25726 Human bre
2	53	25.6	436	5 ABV38691	Abv38691 Human pro
3	53	25.6	440	5 ABV08796	Abv08796 Human pro
4	53	25.6	470	4 AAL18121	Aal18121 Human bre
5	53	25.6	508	9 ACH47938	Ach47938 Human lun
6	49	23.7	468	4 AAL07980	Aal07980 Human bre

7	36	17.4	363	9	ACH42916	Ach42916 Human foe
8	34	16.4	357	9	ACH42960	Ach42960 Human foe
9	30	14.5	346	8	ABX47578	Abx47578 Bovine ES
10	17	8.2	252	6	ABN18436	Abn18436 Human ORF
11	14	6.8	394	8	ABX45923	Abx45923 Bovine ES
12	14	6.8	430	10	ADP80136	Adp80136 Leukaemia
13	13	6.3	181	5	ADL41517	Adl41517 Human ova
14	13	6.3	262	4	ADL14839	Adl14839 Human PSI
15	13	6.3	425	5	ADL176281	Adl176281 Human ova
16	13	6.3	425	5	ADI69948	Adi69948 Human ova
17	13	6.3	439	4	AAI70197	Aai70197 Human pla
18	13	6.3	468	8	ABX45686	Abx45686 Bovine ES
19	13	6.3	479	5	ADI75815	Adi75815 Human ova
20	13	6.3	479	5	ADI69475	Adi69475 Human ova
21	13	6.3	498	5	ADL41051	Adl41051 Human ova
22	13	6.3	546	10	ADE85112	Ades85112 Farnesyl
23	13	6.3	614	6	AAD27741	Aad27741 Human ful
24	12	5.8	304	4	AAK53546	Aak53546 Murine tr
25	12	5.8	447	3	AAC93926	Aac93926 Cat flea
26	12	5.8	543	12	ADL11841	Adl11841 Cat flea
27	12	5.8	549	12	ADL11672	Adl11672 Cat flea
28	12	5.8	549	12	ADL11846	Adl11846 Cat flea
29	12	5.8	550	12	ADL11728	Adl11728 Cat flea
30	11	5.3	113	10	ADG42726	Adg42726 Human CDN
31	11	5.3	113	10	ADJ55795	Adj55795 Fragment
32	11	5.3	113	12	ADM76633	Adm76633 Human NOV
33	11	5.3	205	2	AAQ31930	Aaq31930 Human scu
34	11	5.3	249	8	ABX46694	Abx46694 Bovine ES
35	11	5.3	249	8	ABX36245	Abx36245 Bovine ES
36	11	5.3	252	2	AAT40853	Aat40853 Serine pr
37	11	5.3	252	4	AAV04593	Aav04593 Flea seri
38	11	5.3	252	4	AAC90861	Aac90861 Flea seri
39	11	5.3	262	8	ABX42926	Abx42926 Bovine ES
40	11	5.3	296	6	AAI16931	Aai16931 Human epi
41	11	5.3	307	8	ADA05695	Ada05695 Human NOV
42	11	5.3	307	12	ADN62859	Adn62859 Human NOV
43	11	5.3	321	6	ABN25252	Abn25252 Human ORF
44	11	5.3	358	8	ADA05693	Ada05693 Human NOV
45	11	5.3	358	12	ADN62857	Adn62857 Human NOV

ALIGNMENTS

RESULT 1

AAL25726	
ID	AAL25726 standard; cDNA; 356 BP.
XX	
AC	AAL25726;
XX	
DT	07-DEC-2001 (first entry)
XX	
DE	Human breast cancer expressed polynucleotide 18183.
XX	
KW	Human; breast cancer; cell marker; cytostatic; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200151628-A2.
XX	
PD	19-JUL-2001.
XX	
PF	10-JAN-2001; 2001WO-US000798.
XX	
PR	14-JAN-2000; 2000US-0176077P.
PR	14-MAR-2000; 2000US-0189167P.
PR	24-MAR-2000; 2000US-0192099P.
PR	29-MAR-2000; 2000US-0193480P.
PR	15-MAY-2000; 2000US-0205230P.
PR	09-JUN-2000; 2000US-0211315P.
PR	25-JUL-2000; 2000US-0220534P.
XX	
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;  
XX  
DR WPI; 2001-451856/48.  
XX  
PT New peptide useful as a marker for the diagnosis of breast cancer.  
XX  
PS Claim 1; Page 3360; 3695pp; English.  
XX  
CC The invention relates to human breast cancer expressed polynucleotides  
CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
CC afflicted with breast cancer by examining the correlation between the  
CC expression of certain markers and the cancerous state of breast cells.  
CC The polynucleotides and encoded polypeptides are potential markers for  
CC detecting, diagnosing, monitoring, characterizing treating and  
CC potentially preventing breast cancer. The polynucleotides and encoded  
CC polypeptides are also useful for isolating compounds with cytostatic  
CC activity  
XX  
SQ Sequence 356 BP; 98 A; 71 C; 86 G; 101 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 4,038-42 Length: 356  
Score: 53.00 Matches: 53  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 25.60% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-992-095B-54 (1-207) x AAL25726 (1-356)  
QY 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174  
DB 152 GGTGACAGTGGAGGCGCTCTGGTTTCTTGAGAAAGGACAAATACATTTTACAAGGAGTC 211  
  
QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyrValArgValSer 194  
DB 212 ACTTCTTGGGCTCTGGCTGTGCACGCCCAATAGCTGTGTGTCTATGTTGCTGTTTCA 271  
  
QY 195 ArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207  
DB 272 AGTTTGTACTTGGATTGAGGAGCTGATGAGAAATAAT 310  
  
RESULT 2  
ID ABV38691 standard; cDNA; 436 BP.  
XX  
AC ABV38691;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 38682.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 7870; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 436 BP; 117 A; 87 C; 116 G; 116 T; 0 U; 0 Other;  
  
Alignment Scores:  
Pred. No.: 4,858-42 Length: 436  
Score: 53.00 Matches: 53  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 25.60% Indels: 0  
DB: 5 Gaps: 0  
  
US-09-992-095B-54 (1-207) x ABV38691 (1-436)  
QY 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyVal 174  
DB 199 GGTGACAGTGGAGGCGCTCTGGTTTCTTGAGAAAGGACAAATACATTTTACAAGGAGTC 258  
  
QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyrValArgValSer 194  
DB 259 ACTTCTTGGGCTCTGGCTGTGCACGCCCAATAGCTGTGTGTCTATGTTGCTGTTTCA 318  
  
RESULT 3  
ID ABV08796  
XX  
AC ABV08796 standard; cDNA; 440 BP.  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 8787.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA Schiegel R, Endege WO, Monahan JE;  
 PI WPI; 2001-662795/76.  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 XX prostate cells and correlating with presence of prostate cancer. useful  
 XX for detecting presence of prostate cancer, stage of prostate cancer.  
 XX Claim 1; Page 1386; 11750pp; English.  
 XX The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX Sequence 440 BP; 115 A; 88 C; 120 G; 116 T; 0 U; 1 Other;

Alignment Scores:  
 Pred. No.: 4,89e-42 Length: 440  
 Score: 53.00 Matches: 53  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 25.60% Indels: 0  
 DB: 5 Gaps: 0

US-09-992-095B-54 (1-207) x ABV08796 (1-440)

QY 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrlleLeuGlnGlyVal 174  
 DB 198 GGTGACAGTGGAGGGCTCTGGTTGCTTTGAGAGGACAAATACATTTTACAAGGAGTC 257  
 QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTrValArgValSer 194  
 DB 258 ACTTCTTGGGGCTTGGCTGTGACGCCCAATAGCCCTGGTGTCTATGTTGTTTCA 317  
 QY 195 ArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207  
 DB 318 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 356

RESULT 4  
 AAL18121  
 ID AAL18121 standard; cDNA; 470 BP.  
 XX AAL18121;  
 XX 07-DEC-2001 (first entry)  
 XX Human breast cancer expressed polynucleotide 10578.  
 XX Human; breast cancer; cell marker; cytostatic; ss.  
 XX Homo sapiens.  
 OS WO200151628-A2.  
 PN 19-JUL-2001.  
 PD 10-JAN-2001; 2001WO-US000798.  
 PF 14-JAN-2000; 2000US-0176077P.  
 PR 14-MAR-2000; 2000US-0189167P.  
 PR 24-MAR-2000; 2000US-0192099P.

PR 29-MAR-2000; 2000US-0193480P.  
 PR 15-MAY-2000; 2000US-0205230P.  
 PR 09-JUN-2000; 2000US-0211315P.  
 PR 25-JUL-2000; 2000US-0220534P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 PA Lillie J, Xu Y, Wang Y, Steinmann K;  
 PI WPI; 2001-451856/48.  
 XX New peptide useful as a marker for the diagnosis of breast cancer.  
 XX Claim 1; Page 1890; 3695pp; English.  
 XX The invention relates to human breast cancer expressed polynucleotides  
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterizing treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity  
 XX Sequence 470 BP; 131 A; 95 C; 114 G; 130 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 5,19e-42 Length: 470  
 Score: 53.00 Matches: 53  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 25.60% Indels: 0  
 DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x AAL18121 (1-470)

QY 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrlleLeuGlnGlyVal 174  
 DB 156 GGTGACAGTGGAGGGCTCTGGTTGCTTTGAGAGGACAAATACATTTTACAAGGAGTC 215  
 QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTrValArgValSer 194  
 DB 216 ACTTCTTGGGGCTTGGCTGTGACGCCCAATAGCCCTGGTGTCTATGTTGTTTCA 275  
 QY 195 ArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207  
 DB 276 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 314

RESULT 5  
 ACH47938/c  
 ID ACH47938 standard; cDNA; 508 BP.  
 XX ACH47938;  
 XX 13-OCT-2003 (first entry)  
 XX Human lung tumour cDNA #71.  
 XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
 KW genome mapping; biodiversity; genetic disorder.  
 OS Homo sapiens.  
 XX US2003073623-A1.  
 PN 17-APR-2003.  
 PD 30-JUL-2001; 2001US-00918995.  
 PF 30-JUL-2001; 2001US-00918995.  
 PR (DRMA/) DRMANAC R T.



```
XX 17-APR-2003.
XX
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 30128; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX SQ Sequence 363 BP; 81 A; 91 C; 105 G; 85 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 1.29e-25 Length: 363
Score: 36.00 Matches: 36
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.39% Indels: 0
DB: Gaps: 0

US-09-992-095b-54 (1-207) x ACH42916 (1-363)
QY 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20
DB 255 ATGCACCTTCGTGGAGGCACCTTGATATATCCAGAGTGGTGTGACTGCTGCCACTGC 314

QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAla 36
DB 315 TTGGAGAGTCCCAAGGCCTTCATCCTACAAGTCATCCTCGGTGCC 362

RESULT 8
ACH42960
ID ACH42960 standard; cDNA; 357 BP.
XX
XX ACH42960;
XX
XX 13-OCT-2003 (first entry)
XX
XX Human foetal liver/spleen cDNA #176.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX
```

```
KW genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.
XX
XX Claim 1; SEQ ID NO 30172; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for
XX identifying expressed genes or for physical mapping of the human genome,
XX in forensics, in assessing biodiversity, or in identifying mutations
XX responsible for genetic disorders and other traits. The nucleotide
XX sequences are also useful as hybridisation probes, as oligomers for PCR,
XX for chromosome and gene mapping, in the recombinant production of
XX protein, or in generating antisense DNA or RNA. The purified polypeptide
XX is useful for generating antibodies specific for it. The present sequence
XX is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
XX SQ Sequence 357 BP; 80 A; 89 C; 104 G; 84 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.11e-23 Length: 357
Score: 34.00 Matches: 34
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.43% Indels: 0
DB: Gaps: 0

US-09-992-095b-54 (1-207) x ACH42960 (1-357)
QY 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20
DB 255 ATGCACCTTCGTGGAGGCACCTTGATATATCCAGAGTGGTGTGACTGCTGCCACTGC 314

QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeu 34
DB 315 TTGGAGAGTCCCAAGGCCTTCATCCTACAAGTCATCCTG 356

RESULT 9
ABX47578
ID ABX47578 standard; cDNA; 346 BP.
XX
XX ABX47578;
XX
```

DT 21-FEB-2003 (first entry)  
 XX Bovine EST associated with lactation/muscle/fat deposition #12743.  
 DE  
 XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
 KW muscle deposition; fat deposition; genome mapping; gene identification;  
 KW gene analysis; cattle breeding.  
 XX  
 OS Bos Taurus.  
 XX  
 XX US2002137139-A1.  
 PN  
 XX 26-SEP-2002.  
 PD  
 XX 24-SEP-2001; 2001US-00960352.  
 XX  
 XX 12-JAN-1999; 98US-0115707P.  
 PR  
 XX 11-JAN-2000; 2000US-00480902.  
 XX  
 XX (BYATT/) BYATT J C.  
 PA (MATH/) MATHIALAGAN N.  
 PA (TAON/) TAO N.  
 PA (WARR/) WARREN W C.  
 XX  
 XX Byatt JC, Mathialagan N, Tao N, Warren WC;  
 PI WPI; 2003-110599/10.  
 XX  
 XX New nucleic acid associated with lactation, and muscle and fat  
 PT deposition, useful for genome mapping, gene identification and analysis,  
 PT cattle breeding, or for genetically improving cattle.  
 XX  
 PS Claim 2; SEQ ID NO 12743; 245pp; English.  
 XX  
 CC The invention relates to a purified nucleic acid molecule associated with  
 CC lactation or muscle and fat deposition (designated LMFD), derived from  
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second  
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,  
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are  
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
 CC acid linked to a promoter and a 3' non-translated sequence that  
 CC functions in the cell to cause termination of transcription and addition  
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
 CC (2) determining a level or pattern of a molecule in a bovine cell or  
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any  
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a  
 CC complementary nucleic acid molecule obtained from the bovine cell or  
 CC tissue, where hybridisation between the marker nucleic acid and the  
 CC complementary nucleic acid permits the detection of the molecule; and (b)  
 CC detecting the level or pattern of the complementary nucleic acid, where  
 CC the detection of the complementary nucleic acid is predictive of the  
 CC level or pattern of the molecule. The LMFD nucleic acid is used for  
 CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 CC It is useful for genome mapping, gene identification and analysis, cattle  
 CC breeding, preparation of constructs for use in cattle gene expression, or  
 CC for genetically improving cattle. The present sequence is one of the  
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The  
 CC present sequence was not shown in the specification but was obtained in  
 CC electronic format from the USPTO web site:  
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139  
 XX  
 SQ Sequence 346 BP; 79 A; 88 C; 107 G; 72 T; 0 U; 0 Other;  
  
 Alignment Scores:  
 Pred. No.: 8,19e-20 Length: 346  
 Score: 30.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 14.49% Indels: 0  
 DB: 8 Gaps: 0  
  
 US-09-992-095B-54 (1-207) x ABX47578 (1-346)

QY 148 GlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLysAsp 167  
 DB 218 GGAGGGAAGTACAGCTGCCAGGGTGACAGCGGGGGCTCTGGTCTGCTTTGAGAGGAC 277  
 QY 168 LysTyrIleLeuGlnGlyValThrSerTrp 177  
 DB 278 AAATACATCTGCAAGGAGTCACTTCTTGG 307  
  
 RESULT 10  
 ID ABN18436 standard; cDNA; 252 BP.  
 XX  
 AC ABN18436;  
 XX  
 XX 24-JUN-2002 (first entry)  
 XX Human OREFX polynucleotide sequence SEQ ID NO:5349.  
 XX Human; open reading frame; OREFX; gene therapy; cancer; cirrhosis;  
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
 KW hypertension; hypothyroidism; cholesterol ester storage disease;  
 KW immune deficiency; immune disorder; infectious disease;  
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
 KW myasthenia gravis; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200192523-A2.  
 PN  
 XX 06-DEC-2001.  
 PD  
 XX 29-MAY-2001; 2001WO-US010836.  
 XX  
 XX 30-MAY-2000; 2000US-0206132P.  
 PR  
 XX 29-AUG-2000; 2000US-0228716P.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 PA  
 XX Shimkets RA, Leach MD;  
 PI WPI; 2002-106308/14.  
 XX  
 DR P-PSDB; ABP02684.  
 DR  
 XX Novel human polypeptides and polynucleotides useful for diagnosing,  
 PT preventing and treating cardiovascular disease, neurodegenerative,  
 PT hyperproliferative disorders and autoimmune disorders.  
 XX  
 PS Disclosure; SEQ ID NO 5349; 1037pp; English.  
 XX  
 CC The present invention describes substantially purified human proteins  
 CC (referred to as open reading frame, OREFX, where X is 1-11491 (see Table 1  
 CC in the specification). ABN15762 to ABN27252 encode the human OREFX  
 CC proteins given in ABP00010 to ABP11500. OREFX proteins are useful for  
 CC treating or preventing a pathology associated with an OREFX-associated  
 CC disorder in humans, and in the manufacture of a medicament for treating a  
 CC syndrome associated with OREFX-associated disorder. OREFX polynucleotide  
 CC sequences can be used in gene therapy. OREFX sequences can be used in the  
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
 CC osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic  
 CC transplantation, cardiovascular diseases, diabetes related to organ  
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
 CC storage disease, various immune deficiencies and disorders, infectious  
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
 CC disease and autoimmune inflammatory eye disease. OREFX proteins are also  
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
 CC bone degenerative disorders, or periodontal disease, and for gut  
 CC protection or regeneration and treatment of lung or liver fibrosis,  
 CC reperfusion injury in various tissues and conditions resulting from  
 CC systemic cytokine damage. N.B. The sequence data for this patent did not



CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at [fp.wipo.int/pub/published\\_pct\\_sequences](http://fp.wipo.int/pub/published_pct_sequences)  
 XX  
 SQ Sequence 252 BP; 69 A; 70 C; 62 G; 51 T; 0 U; 0 Other;

Alignment Scores: 2.53e-07 Length: 252  
 Pred. No.: 17.00 Matches: 17  
 Score: 100.00% Conservativity: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 8.21% Gaps: 0  
 DB:

US-09-992-095B-54 (1-207) x ABN18436 (1-252)

QY 78 ValileProAlaCysLeuProSerProAenTyrValValAlaAspArgThr 94  
 Db 151 GTCATCCAGCTTGCTGCGCATCTCCAAATTATGTGGTTCGCGACGACA 201

RESULT 11  
 ABX45923  
 ID ABX45923 standard; cDNA; 394 BP.

XX AC ABX45923;  
 XX  
 DT 21-FEB-2003 (first entry)  
 XX  
 DE Bovine EST associated with lactation/muscle/fat deposition #11088.  
 XX  
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;  
 KW muscle deposition; fat deposition; genome mapping; gene identification;  
 KW gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137139-A1.

XX 26-SEP-2002.

XX 24-SEP-2001; 2001US-00960352.

XX 12-JAN-1999; 99US-0115707P.

XX 11-JAN-2000; 2000US-00480902.

XX (BYAT/) BYATT J C.

XX (NATH/) NATHALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-110599/10.  
 XX New nucleic acid associated with lactation, and muscle and fat  
 XX deposition, useful for genome mapping, gene identification and analysis,  
 XX cattle breeding, or for genetically improving cattle.

XX Claim 2; SEQ ID NO 11088; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with  
 XX lactation or muscle and fat deposition (designated LMFD), derived from  
 XX cattle, and the LMFD nucleic acid can specifically hybridise to a second  
 XX nucleic acid molecule comprising any of 1512 nucleotide sequences,  
 XX appearing as ABX34836-ABX49947, or complements of them. Also included are  
 XX ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic  
 XX acid linked to a promoter and a 3' non-translated sequence that  
 XX functions in the cell to cause termination of transcription and addition  
 XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and  
 XX (2) determining a level or pattern of a molecule in a bovine cell or  
 XX tissue comprising: (a) incubating a marker nucleic acid (comprising any  
 XX of the 1512 nucleic acid sequences or its complement or fragment) with a  
 XX complementary nucleic acid molecule obtained from the bovine cell or  
 XX tissue, where hybridisation between the marker nucleic acid and the

CC complementary nucleic acid permits the detection of the molecule; and (b)  
 CC detecting the level or pattern of the complementary nucleic acid, where  
 CC the detection of the complementary nucleic acid is predictive of the  
 CC level or pattern of the molecule. The LMFD nucleic acid is used for  
 CC determining a level or pattern of a molecule in a bovine cell or tissue.  
 CC It is useful for genome mapping, gene identification and analysis, cattle  
 CC breeding, preparation of constructs for use in cattle gene expression, or  
 CC for genetically improving cattle. The present sequence is one of the  
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The  
 CC present sequence was not shown in the specification but was obtained in  
 CC electronic format from the USPTO web site:  
 CC [seqdata.uspto.gov/sequence.html?DocID=20020137139](http://seqdata.uspto.gov/sequence.html?DocID=20020137139)  
 XX

SQ Sequence 394 BP; 100 A; 104 C; 119 G; 71 T; 0 U; 0 Other;

Alignment Scores: 0.000309 Length: 394  
 Pred. No.: 14.00 Matches: 14  
 Score: 100.00% Conservativity: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 6.76% Gaps: 0  
 DB:

US-09-992-095B-54 (1-207) x ABX45923 (1-394)

QY 148 GlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeu 161  
 Db 351 GGAGGGAGTGTGACAGCTGCCAGGTGACAGCGCGGCTCTG 392

RESULT 12

ADP80136

ID ADP80136 standard; DNA; 430 BP.

XX AC ADP80136;

XX 26-FEB-2004 (first entry)

XX Leukaemia-related DNA sequence #692.

XX Cytostatic; Gene therapy; leukaemia; ss.

XX Unidentified.

XX WO2003039443-A2.

XX 15-MAY-2003.

XX 04-NOV-2002; 2002WO-EP012303.

XX 05-NOV-2001; 2001EP-00126244.

XX 30-APR-2002; 2002EP-00009758.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX (UYLU-) UNIV LUDWIG MAXIMILIANS.

XX (HAFE/) HAFERLACH T.

XX (SCHO/) SCHOCH C.

XX (KERN/) KERN W.

XX Haferlach T, Schoch C, Kern W, Kohlmann A, Schnittger S, Dugas M;

XX Bils R, Brors B, Mergenthaler S;

XX WPI; 2003-505037/47.

XX Determining the subtype of leukemia cells and whether a patient sample

XX contains leukemia cells or other cells, useful for treating leukemia,

XX comprises determining the expression profile of a group of markers in a

XX patient sample.

XX Disclosure; SEQ ID NO 692; 2938pp; English.

XX The present invention relates to a method (M1) for determining the

XX subtype of leukaemia cells and whether a patient sample contains

XX leukaemia cells. The method comprises determining the expression profile

CC of a group of markers in a patient sample. The method is useful for  
 CC determining the presence of leukaemia cells, its types or subtypes, and  
 CC for the preparation of a medicament for treating leukaemia.

XX SQ Sequence 430 BP; 103 A; 122 C; 113 G; 92 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 0.000335 Length: 430  
 Score: 14.00 Matches: 14  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.76% Indels: 0  
 DB: 10 Gaps: 0

US-09-992-095B-54 (1-207) x ADF80136 (1-430)

Oy 150 ThrAspSerCysGlnGlyAspSerGlyCysGlyProLeuValCys 163

Db 291 ACAGATTCTCCAGGGAGACTCAGGGGGACCCCTCTGTGT 332

# RESULT 13

ADL41517  
 ID ADL41517 standard; DNA; 181 BP.  
 XX  
 AC ADL41517;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human ovarian cancer DNA marker #15407.  
 XX  
 KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200170979-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 21-MAR-2001; 2001WO-US009136.  
 XX  
 PR 21-MAR-2000; 2000US-0191031P.  
 PR 25-MAY-2000; 2000US-0207124P.  
 PR 15-JUN-2000; 2000US-0211940P.  
 PR 07-JUL-2000; 2000US-0216820P.  
 PR 25-JUL-2000; 2000US-0220661P.  
 PR 21-DEC-2000; 2000US-0257672P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Lee J, Lillie J;  
 XX  
 DR WPI; 2001-611502/70.  
 XX  
 PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian  
 PT cancer cells as compared to their normal non-cancerous ovarian cells are  
 PT used to characterize stage, grade, histological type of ovarian cancer.  
 XX  
 PS Disclosure; SEQ ID NO 15407; 106pp; English.  
 XX

CC The invention relates to nucleic acid markers which are overexpressed in  
 CC ovarian cancer cells as compared to their expression in normal (i.e. non-  
 CC cancerous) ovarian cells. The invention also relates to polypeptides  
 CC encoded by the markers, antibodies that selectively bind to the  
 CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk  
 CC of developing ovarian cancer involving inhibiting expression of a gene  
 CC corresponding to a marker of the invention and a method of treating a  
 CC patient afflicted with ovarian cancer comprising providing to cells of  
 CC the patient an antisense oligonucleotide complementary to a marker of the  
 CC invention. The markers are useful for assessing if a patient is afflicted  
 CC with ovarian cancer, which involves comparing the level of expression of  
 CC a marker in a patient sample and a normal level of expression of the  
 CC marker in a control non-ovarian cancer sample. A difference between the  
 CC expression levels indicates ovarian cancer. The level of expression of a

CC marker corresponds to a secreted protein or to a transcribed  
 CC polynucleotide or its portion. The level of expression of the marker is  
 CC assessed by detecting the presence in the sample, a protein or protein  
 CC fragment corresponding to the marker. The presence of protein or protein  
 CC fragment is detected using an antibody that specifically binds with the  
 CC protein or protein fragment. Alternatively, the level of expression of  
 CC the marker is assessed by detecting the presence of a transcribed  
 CC polynucleotide which anneals with the marker or anneals with a portion of  
 CC the polynucleotide comprising the marker, under stringent conditions. The  
 CC marker is also used for monitoring the progression of ovarian cancer in a  
 CC patient which involves detecting expression of the marker in a patient  
 CC sample at a first point in time, repeating the method at a subsequent  
 CC time and comparing the level of expression. The method is carried out  
 CC using an ovarian tissue sample. A composition comprising a marker,  
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.  
 CC This sequence represents a human ovarian cancer DNA marker of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 181 BP; 38 A; 48 C; 62 G; 33 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 0.00143 Length: 181  
 Score: 13.00 Matches: 13  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.28% Indels: 0  
 DB: 5 Gaps: 0

US-09-992-095B-54 (1-207) x ADL41517 (1-181)

Oy 151 AspSerCysGlnGlyAspSerGlyCysGlyProLeuValCys 163

Db 66 GATTCCTGCCAGGTGATTCTGGGGGTCGCTGTATGT 104

# RESULT 14

AAD14839  
 ID AAD14839 standard; cDNA; 262 BP.  
 XX  
 AC AAD14839;  
 XX  
 DT 01-NOV-2001 (first entry)  
 XX  
 DE Human PS133 cDNA from clone 1856238.  
 XX  
 KW Human; PS133; prostate disease; cancer; immunogen; gene therapy; EST;  
 KW expressed sequence tag; cytostatic; clone 1856238; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6232456-B1.  
 XX  
 PD 15-MAY-2001.  
 XX  
 PF 06-OCT-1997; 97US-00944483.  
 XX  
 PR 06-OCT-1997; 97US-00944483.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Cohen M, Colpitts TL, Friedman PN, Granados E, Klass MR;  
 PI Russell JC, Stewart KD, Stroupe SD;  
 XX  
 DR WPI; 2001-366357/38.  
 XX  
 PT New PS133 polynucleotides, useful for detecting, diagnosing, staging,  
 PT monitoring, prognosing, preventing, treating or determining the  
 PT predisposition of an individual to a prostate disease, e.g. cancer.  
 XX  
 PS Claim 1; Fig 1; 93pp; English.  
 XX  
 CC The patent discloses PS133 polynucleotides and polypeptides which are

CC indicative of prostate disease. The patent also provides a method for  
 CC detecting PS133 protein in a test sample. The polynucleotides of the  
 CC invention are useful for detecting, diagnosing, staging, monitoring,  
 CC prognosing, preventing, treating or determining the predisposition of an  
 CC individual to prostate diseases such as cancer. PS133-derived  
 CC polynucleotides are used for the detection of normal or altered gene  
 CC expression, in assays for detecting, amplifying or quantifying genes or  
 CC nucleic acids relating to prostate tissue diseases and conditions, and to  
 CC produce probes which can be used in the detection of nucleic acids in a  
 CC sample. PS133 proteins are used as immunogens for the production of  
 CC antibodies. PS133 sequences are also used in gene therapy. The present  
 CC sequence is human PS133 cDNA fragment derived from overlapping expressed  
 CC sequence tag (EST) clone 1856238

XX Sequence 262 BP; 65 A; 81 C; 70 G; 46 T; 0 U; 0 Other;

Alignment Scores: Length: 262  
 Pred. No.: 0.00199 Matches: 13  
 Score: 13.00 Conservations: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 6.28% Gaps: 0  
 DB: 4

US-09-992-095B-54 (1-207) x AAD14839 (1-262)

QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 163  
 DB 64 GACTCTGCCAGGGTACTCCGGGGCCCTCTGGTCTGT 102

RESULT 15

AD176281  
 ID AD176281 standard; DNA; 425 BP.

AC AD176281;

DT 20-MAY-2004 (first entry)

XX Human ovarian cancer DNA marker #9023.

XX Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

OS Homo sapiens.

PN W0200170979-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009126.

XX 21-MAR-2000; 2000US-0191031P.

PR 25-MAY-2000; 2000US-0207124P.

PR 15-JUN-2000; 2000US-0211940P.

PR 07-JUL-2000; 2000US-0216820P.

PR 25-JUL-2000; 2000US-0220661P.

PR 21-DEC-2000; 2000US-0257672P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lee J, Lillie J;

XX WPI; 2001-611502/70.

XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian  
 XX cancer cells as compared to their normal non-cancerous ovarian cells are  
 XX used to characterize stage, grade, histological type of ovarian cancer.

XX Disclosure; SEQ ID NO 9023; 106pp; English.

XX The invention relates to nucleic acid markers which are overexpressed in  
 XX ovarian cancer cells as compared to their expression in normal (i.e. non-  
 XX cancerous) ovarian cells. The invention also relates to polypeptides  
 XX encoded by the markers, antibodies that selectively bind to the

CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk  
 CC of developing ovarian cancer involving inhibiting expression of a gene  
 CC corresponding to a marker of the invention and a method of treating a  
 CC patient afflicted with ovarian cancer comprising providing to cells of  
 CC the patient an antisense oligonucleotide complementary to a marker of the  
 CC invention. The markers are useful for assessing if a patient is afflicted  
 CC with ovarian cancer, which involves comparing the level of expression of  
 CC a marker in a patient sample and a normal level of expression of the  
 CC marker in a control non-ovarian cancer sample. A difference between the  
 CC expression levels indicates ovarian cancer. The level of expression of a  
 CC marker corresponds to a secreted protein or to a transcribed  
 CC polynucleotide or its portion. The level of expression of the marker is  
 CC assessed by detecting the presence in the sample, a protein or protein  
 CC fragment corresponding to the marker. The presence of protein or protein  
 CC fragment is detected using an antibody that specifically binds with the  
 CC protein or protein fragment. Alternatively, the level of expression of  
 CC the marker is assessed by detecting the presence of a transcribed  
 CC polynucleotide which anneals with the marker or anneals with a portion of  
 CC the polynucleotide comprising the marker, under stringent conditions. The  
 CC marker is also used for monitoring the progression of ovarian cancer in a  
 CC patient which involves detecting expression of the marker in a patient  
 CC sample at a first point in time, repeating the method at a subsequent  
 CC time and comparing the level of expression. The method is carried out  
 CC using an ovarian tissue sample. A composition comprising a marker, cancer.  
 CC polypeptide or antibody of the invention is used to treat ovarian cancer.  
 CC This sequence represents a human ovarian cancer DNA marker of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
 SQ Sequence 425 BP; 94 A; 111 C; 107 G; 98 T; 0 U; 15 Other;

Alignment Scores: Length: 425  
 Pred. No.: 0.0031 Matches: 13  
 Score: 13.00 Conservations: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 100.00% Indels: 0  
 Query Match: 6.28% Gaps: 0  
 DB: 5

US-09-992-095B-54 (1-207) x AD176281 (1-425)

QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 163  
 DB 66 GATTCTGCCAGGGTACTTCGGGGTCCGCTGGTATGT 104

Search completed: October 27, 2004, 09:48:47  
 Job time : 510 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: October 27, 2004, 09:38:33 ; Search time 106 Seconds  
(without alignments)  
1388.050 Million cell updates/sec

Title: US-09-992-095B-54  
Perfect score: 207  
Sequence: 1 MHFCGTLSPWVLTAAHC.....GVYVRVRFVTWIEGVMRNN 207

Scoring table: OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 824507 seqs, 355394441 residues  
Word size: 1  
Total number of hits satisfying chosen parameters: 1385334

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US0992095/runat\_26102004\_090253\_11831/app.query.fasta\_1.391  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=621  
-USER=US0992095 @CGN\_1\_1\_69 brunat\_26102004\_090253\_11831 -NCPUS=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	13.0	312	4	US-09-621-976-9524
2	13	6.3	262	3	US-08-944-483-5
3	13	6.3	614	4	US-09-879-792-33
4	11	5.3	252	3	US-08-906-769-98
5	11	5.3	252	3	US-08-906-616-98
6	11	5.3	252	3	US-08-817-795-98
7	11	5.3	252	3	US-08-639-075A-98
8	11	5.3	252	3	US-09-012-431-98
9	11	5.3	252	3	US-09-012-692-98
10	11	5.3	252	3	US-08-906-613-98
11	11	5.3	252	5	PCT-US95-14442A-98
12	11	5.3	290	4	US-09-600-985-4

13	11	5.3	290	4	US-09-600-985-5
14	11	5.3	290	4	US-09-600-985-6
15	11	5.3	290	4	US-09-600-985-7
16	11	5.3	453	6	5200340-3
17	11	5.3	506	3	US-09-280-116-72
18	11	4.8	340	3	US-09-280-116-28
19	10	4.8	340	3	US-09-280-116-29
20	9	4.3	27	1	US-08-485-455D-11
21	9	4.3	27	2	US-08-482-130C-11
22	9	4.3	27	2	US-08-484-211C-11
23	9	4.3	27	3	US-08-906-769-11
24	9	4.3	27	3	US-08-906-616-11
25	9	4.3	27	3	US-08-817-795-11
26	9	4.3	27	3	US-08-485-443B-11
27	9	4.3	27	3	US-08-639-075A-11
28	9	4.3	27	3	US-09-012-431-11
29	9	4.3	27	3	US-09-032-215-17
30	9	4.3	27	3	US-09-012-692-11
31	9	4.3	27	3	US-08-906-613-11
32	9	4.3	27	5	PCT-US95-14442A-11
33	9	4.3	60	2	US-08-832-468-4
34	9	4.3	155	1	US-08-518-878B-25
35	9	4.3	155	1	US-08-518-878B-26
36	9	4.3	155	1	US-08-294-522B-25
37	9	4.3	155	1	US-08-294-522B-26
38	9	4.3	155	2	US-08-807-861A-25
39	9	4.3	155	2	US-08-807-861A-26
40	9	4.3	155	2	US-08-470-868A-25
41	9	4.3	155	2	US-08-470-868A-26
42	9	4.3	155	3	US-09-210-681-25
43	9	4.3	155	3	US-09-210-681-26
44	9	4.3	155	3	US-08-946-719A-25
45	9	4.3	155	3	US-08-946-719A-26

ALIGNMENTS

RESULT 1  
US-09-621-976-9524  
; Sequence 9524, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 9524  
; LENGTH: 312  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 189  
; OTHER INFORMATION: n=a, g, c o r t  
US-09-621-976-9524

Alignment Scores:  
Pred. No.: 1.31e-20 Length: 312  
Score: 27.00 Matches: 27  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 13.04% Indels: 0  
DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-621-976-9524 (1-312)

QY 122 GluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeu 141  
|||||

Db 2 GAGATTAAGTGTCAATCGCTATGAGTTTCTGAATGGAAGAGTCCATCCACCGACTC 61

Qy 142 CysAlaGlyHisLeuAlaGly 148  
|||||

Db 62 TGTGCTGGCATTGTCGGCGA 82

RESULT 2

US-08-944-483-5  
; Sequence 5, Application US/08944483  
; Patent No. 6232456  
; GENERAL INFORMATION:  
; APPLICANT: COHEN, MAURICE  
; APPLICANT: COLPITTS, TRACEY L.  
; APPLICANT: FRIEDMAN, PAULA N.  
; APPLICANT: GRANADOS, EDWARD N.  
; APPLICANT: KLASS, MICHAEL R.  
; APPLICANT: RUSSELL, JOHN C.  
; APPLICANT: STEWART, KENT D.  
; APPLICANT: STROUPE, STEVEN D.  
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS  
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES  
; NUMBER OF SEQUENCES: 76  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-SEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/944,483  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Becker, Cheryl L.  
; REGISTRATION NUMBER: 35,441  
; REFERENCE/DOCKET NUMBER: 6183.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 847/935-1729  
; TELEFAX: 847/938-2623  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 262 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-944-483-5

Alignment Scores:  
Pred. No.: 2,448-05 Length: 262  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-944-483-5 (1-262)

Qy 151 AspSerCysGlnGlyAspSerGlyProLeuValCys 163  
|||||

Db 64 GACTCTGCAGGGTGACTCCGGGGGCCCTCTGCTGT 102  
|||||

RESULT 3

US-09-879-792-33  
; Sequence 33, Application US/09879792  
; Patent No. 6734006  
; GENERAL INFORMATION:  
; APPLICANT: Xiao, Yonghong  
; APPLICANT: Gedrich, Richard  
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine  
; TITLE OF INVENTION: Protease  
; FILE REFERENCE: 02973.00035  
; CURRENT APPLICATION NUMBER: US/09/879,792  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/211,224  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US 60/283,353  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/283,648  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Fast-SEQ for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 614  
; TYPE: DNA  
; ORGANISM: mouse  
US-09-879-792-33

Alignment Scores:  
Pred. No.: 5,81e-05 Length: 614  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-879-792-33 (1-614)

Qy 151 AspSerCysGlnGlyAspSerGlyProLeuValCys 163  
|||||

Db 318 GACTCTGCCAGGAGACAGTGGAGACCTCTCGTGT 356  
|||||

RESULT 4

US-08-906-769-98  
; Sequence 98, Application US/08906769  
; Patent No. 6077687  
; GENERAL INFORMATION:  
; APPLICANT: Grieve, Robert B.  
; APPLICANT: Rushlow, Keith E.  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Frank, Glenn R.  
; APPLICANT: Stiegler, Gary  
; APPLICANT: Gaines, Patrick J.  
; APPLICANT: Silver, Gary  
; TITLE OF INVENTION: PLEA PROTEASE PROTEINS, NUCLEIC ACID  
; TITLE OF INVENTION: MOLECULES AND USES THEREOF  
; NUMBER OF SEQUENCES: 190  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/906,769  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/639,075

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-25-C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 98:

SEQUENCE CHARACTERISTICS:

LENGTH: 252 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..252

US-08-906-769-98

Alignment Scores:

Pred. No.: 0.00365

Score: 11.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 5.31%

DB: 3

Length: 252

Matches: 11

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-992-095B-54 (1-207) x US-08-906-769-98 (1-252)

QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161

DB 220 GATTCTTGTCAGGAGACTCTGGTGGCTCTTTA 252

RESULT 5

US-08-906-616-98

Sequence 98, Application US/08906616

Patent No. 6121035

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.

APPLICANT: Rushlow, Keith E.

APPLICANT: Wu Hunter, Shirley

APPLICANT: Frank, Glenn R.

APPLICANT: Stiegler, Gary

APPLICANT: Gaines, Patrick J.

APPLICANT: Silver, Gary

TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF

NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross P.C.

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/906,616

FILING DATE: 05-AUG-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-25-C2-3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 98:

SEQUENCE CHARACTERISTICS:

LENGTH: 252 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..252

US-08-906-616-98

Alignment Scores:

Pred. No.: 0.00365

Score: 11.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 5.31%

DB: 3

Length: 252

Matches: 11

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-992-095B-54 (1-207) x US-08-906-616-98 (1-252)

QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161

DB 220 GATTCTTGTCAGGAGACTCTGGTGGCTCTTTA 252

RESULT 6

US-08-817-795-98

Sequence 98, Application US/08817795

Patent No. 6139840

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.

APPLICANT: Rushlow, Keith E.

APPLICANT: Hunter, Shirley Wu

APPLICANT: Frank, Glenn R.

APPLICANT: Heath, Andrew W.

APPLICANT: Yamaka, Miles Yamanaka

APPLICANT: Arfsten, Ann

APPLICANT: Stiegler, Gary

TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND

TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA

TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID

TITLE OF INVENTION: MOLECULES, AND USES THEREOF

NUMBER OF SEQUENCES: 119

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/817,795

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/14442

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Gary J. Connell

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 98:

SEQUENCE CHARACTERISTICS:

LENGTH: 252 base pairs

TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..252  
US-08-817-795-98

Alignment Scores:  
Pred. No.: 0.00365 Length: 252  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-817-795-98 (1-252)

Qy 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161  
Db 220 GATTCTTGTCAAGGAGACTCTGGTGGTCCTTTA 252

RESULT 7  
US-08-639-075A-98  
Sequence 98, Application US/08639075A  
Patent No. 6150125  
GENERAL INFORMATION:  
APPLICANT: Grieve, Robert B.  
APPLICANT: Rushlow, Keith E.  
APPLICANT: Wu Hunter, Shirley  
APPLICANT: Frank, Glenn R.  
APPLICANT: Stiegler, Gary  
APPLICANT: Gaines, Patrick J.  
APPLICANT: Silver, Gary  
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
TITLE OF INVENTION: MOLECULES AND USES THEREOF  
NUMBER OF SEQUENCES: 190  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/639,075A  
FILING DATE: 24-APR-1996  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-25-C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..252  
US-08-639-075A-98

Alignment Scores:  
Pred. No.: 0.00365 Length: 252  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-639-075A-98 (1-252)

Qy 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161  
Db 220 GATTCTTGTCAAGGAGACTCTGGTGGTCCTTTA 252

RESULT 8  
US-09-012-431-98  
Sequence 98, Application US/09012431  
Patent No. 6180383  
GENERAL INFORMATION:  
APPLICANT: Grieve, Robert B.  
APPLICANT: Rushlow, Keith E.  
APPLICANT: Wu Hunter, Shirley  
APPLICANT: Frank, Glenn R.  
APPLICANT: Stiegler, Gary  
APPLICANT: Gaines, Patrick J.  
APPLICANT: Silver, Gary  
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
TITLE OF INVENTION: MOLECULES AND USES THEREOF  
NUMBER OF SEQUENCES: 190  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/012,431  
FILING DATE: 23-Jan-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/639,075  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-25-C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 252 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..252  
US-09-012-431-98

Alignment Scores:  
Pred. No.: 0.00365 Length: 252  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0



Query Match: 5.31% Indels: 0  
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-012-431-98 (1-252)

QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161  
|||||  
DB 220 GATTCTTGTCAAGGAGACTCTGTGTCCTTTA 252

RESULT 9

US-09-012-692-98  
; Sequence 98, Application US/09012692  
; Patent No. 6214579  
; GENERAL INFORMATION:  
; APPLICANT: Grieve, Robert B.  
; APPLICANT: Rushlow, Keith E.  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Frank, Glenn R.  
; APPLICANT: Stiegler, Gary  
; APPLICANT: Gaines, Patrick J.  
; APPLICANT: Silver, Gary  
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
; TITLE OF INVENTION: MOLECULES AND USES THEREOF  
; NUMBER OF SEQUENCES: 190  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/012,692  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/639,075  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-25-C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 252 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..252

US-09-012-692-98

Alignment Scores:  
Pred. No.: 0.00365 Length: 252  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-012-692-98 (1-252)

QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161

DB 220 GATTCTTGTCAAGGAGACTCTGTGTCCTTTA 252

RESULT 10

US-08-906-613-98  
; Sequence 98, Application US/08906613  
; Patent No. 6232096  
; GENERAL INFORMATION:  
; APPLICANT: Grieve, Robert B.  
; APPLICANT: Rushlow, Keith E.  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Frank, Glenn R.  
; APPLICANT: Stiegler, Gary  
; APPLICANT: Gaines, Patrick J.  
; APPLICANT: Silver, Gary  
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
; TITLE OF INVENTION: MOLECULES AND USES THEREOF  
; NUMBER OF SEQUENCES: 190  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/906,613  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/639,075  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-25-C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 252 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..252

US-08-906-613-98

Alignment Scores:  
Pred. No.: 0.00365 Length: 252  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 3 Gaps: 0

US-09-992-095B-54 (1-207) x US-08-906-613-98 (1-252)

QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161  
|||||  
DB 220 GATTCTTGTCAAGGAGACTCTGTGTCCTTTA 252

RESULT 11

PCT-US95-14442A-98  
; Sequence 98, Application PC/TUS9514442A

```
;
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Frank, Glenn R.
; APPLICANT: Heath, Andrew W.
; APPLICANT: Yamaka, Miles Yamanaka
; APPLICANT: Arfsten, Ann
; APPLICANT: Dale, Beverly
; APPLICANT: Stiegler, Gary
; TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND
; TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA
; TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14442A
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gary J. Connell
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..252
; PCT-US95-14442A-98

Alignment Scores:
Pred. No.: 0.00365 Length: 252
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.31% Indels: 0
DB: 5 Gaps: 0

US-09-992-095B-54 (1-207) x PCT-US95-14442A-98 (1-252)
QY 151 AspSerCysGlnGlyAspSerGlyGlyProLeu 161
DB 220 GATTCTGTGCAAGGAGACTCTGCTGCTCTTGA 252

RESULT 12
US-09-600-985-4
; Sequence 4, Application US/09600985
; Patent No. 6706504
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin L
; TITLE OF INVENTION: TISSUE TYPE PLASMINOGEN ACTIVATOR (t-PA)
; TITLE OF INVENTION: VARIANTS HAVING ZYMOMEN CHARACTERISTICS: COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: TSRI 568.1
; CURRENT APPLICATION NUMBER: US/09/600,985
; CURRENT FILING DATE: 2000-11-13
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 290
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-600-985-5
; Sequence 5, Application US/09600985
; Patent No. 6706504
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin L
; TITLE OF INVENTION: TISSUE TYPE PLASMINOGEN ACTIVATOR (t-PA)
; TITLE OF INVENTION: VARIANTS HAVING ZYMOMEN CHARACTERISTICS: COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: TSRI 568.1
; CURRENT APPLICATION NUMBER: US/09/600,985
; CURRENT FILING DATE: 2000-11-13
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 290
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-600-985-5
Alignment Scores:
Pred. No.: 0.00422 Length: 290
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.31% Indels: 0
DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-600-985-5 (1-290)
QY 153 CysGlnGlyAspSerGlyGlyProLeuValCys 163
DB 182 TGCAGGGCGGATTCGGAGGAGCCCTGCTGTGT 214

RESULT 13
US-09-600-985-5
; Sequence 5, Application US/09600985
; Patent No. 6706504
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin L
; TITLE OF INVENTION: TISSUE TYPE PLASMINOGEN ACTIVATOR (t-PA)
; TITLE OF INVENTION: VARIANTS HAVING ZYMOMEN CHARACTERISTICS: COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: TSRI 568.1
; CURRENT APPLICATION NUMBER: US/09/600,985
; CURRENT FILING DATE: 2000-11-13
; PRIOR FILING DATE: 1996-11-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 290
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-600-985-5
Alignment Scores:
Pred. No.: 0.00422 Length: 290
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.31% Indels: 0
DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-600-985-5 (1-290)
QY 153 CysGlnGlyAspSerGlyGlyProLeuValCys 163
DB 182 TGCAGGGCGGATTCGGAGGAGCCCTGCTGTGT 214

RESULT 14
US-09-600-985-6
; Sequence 6, Application US/09600985
; Patent No. 6706504
; GENERAL INFORMATION:
; APPLICANT: Madison, Edwin L
; TITLE OF INVENTION: TISSUE TYPE PLASMINOGEN ACTIVATOR (t-PA)
; TITLE OF INVENTION: VARIANTS HAVING ZYMOMEN CHARACTERISTICS: COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: TSRI 568.1
```

; CURRENT APPLICATION NUMBER: US/09/600,985  
 ; CURRENT FILING DATE: 2000-11-13  
 ; PRIOR APPLICATION NUMBER: PCT/WO98/21320  
 ; PRIOR FILING DATE: 1996-11-12  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6  
 ; LENGTH: 290  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-600-985-6

Alignment Scores:  
 Pred. No.: 0.00422 Length: 290  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.31% Indels: 0  
 DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-600-985-6 (1-290)

QY 153 CysGlnGlyAspSerGlyProLeuValCys 163  
 Db 182 TCCAGGGCGATTCCGGAGGCCCCCTGGTGT 214

RESULT 15

US-09-600-985-7  
 ; Sequence 7, Application US/09600985  
 ; Patent No. 6706504  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Madison, Edwin L.  
 ; TITLE OF INVENTION: TISSUE TYPE PLASMINOGEN ACTIVATOR (t-PA)  
 ; TITLE OF INVENTION: VARIANTS HAVING ZYMOMEN CHARACTERISTICS: COMPOSITIONS AND  
 ; TITLE OF INVENTION: METHODS OF USE  
 ; FILE REFERENCE: TSRI 568.1  
 ; CURRENT APPLICATION NUMBER: US/09/600,985  
 ; CURRENT FILING DATE: 2000-11-13  
 ; PRIOR APPLICATION NUMBER: PCT/WO98/21320  
 ; PRIOR FILING DATE: 1996-11-12  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 290  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-600-985-7

Alignment Scores:  
 Pred. No.: 0.00422 Length: 290  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.31% Indels: 0  
 DB: 4 Gaps: 0

US-09-992-095B-54 (1-207) x US-09-600-985-7 (1-290)

QY 153 CysGlnGlyAspSerGlyProLeuValCys 163  
 Db 182 TCCAGGGCGATTCCGGAGGCCCCCTGGTGT 214

Search completed: October 27, 2004, 11:49:49  
 Job time : 108 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 27, 2004, 10:56:55 ; Search time 493 Seconds

(without alignments)  
2152.969 Million cell updates/sec

Title: US-09-992-095B-54

Perfect score: 207

Sequence: 1 MHFCGTLSPWLVTAHC.....GVYVRVRFVTWIEGVRNN 207

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3413475 seqs, 2563800928 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4661235

Minimum DB seq length: 0

Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US0992095/runat\_26102004\_090255\_11917/app\_query.fasta\_1.391  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=621 -USER=US0992095@cgn 1\_1\_480 @runat\_26102004\_090255\_11917 -NCPUS=6  
-ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -XGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Alignment Scores: 1.65e-60 Length: 201  
Pred. No.: 201

Result No.	Score	Query Match	Length	DB ID	Description
1	66	31.9	201	17	US-10-741-601-2633
2	66	31.9	201	17	US-10-741-601-2649
3	66	31.9	201	17	US-10-741-601-2657
4	66	31.9	201	17	US-10-741-601-2660
5	53	25.6	201	17	US-10-741-601-2630
6	53	25.6	201	17	US-10-741-601-14280
7	53	25.6	508	10	US-09-918-995-35150
8	46	22.2	201	17	US-10-741-601-14344
9	38	18.4	201	17	US-10-741-601-14315
10	37	17.9	201	17	US-10-741-601-14376
11	37	17.9	201	17	US-10-741-601-14492
12	36	17.4	363	10	US-09-918-995-30128
13	34	16.4	357	10	US-09-918-995-30172
14	33	15.9	201	17	US-10-741-601-2639
15	33	15.9	201	17	US-10-741-601-2653
16	32	15.5	201	17	US-10-741-601-2629
17	30	14.5	346	9	US-09-960-352-12743
18	27	13.0	201	17	US-10-741-601-14453
19	24	11.6	201	17	US-10-741-601-14495
20	22	10.6	201	17	US-10-741-601-14401
21	21	10.1	201	17	US-10-741-601-14486
22	17	8.2	201	17	US-10-741-601-2659
23	17	8.2	201	17	US-10-741-601-14374
24	16	7.7	201	17	US-10-741-601-14327
25	16	7.7	201	17	US-10-741-601-14445
26	14	6.8	201	17	US-10-741-601-14420
27	14	6.8	379	9	US-09-796-692-4052
28	14	6.8	379	14	US-10-040-862-4052
29	14	6.8	379	16	US-10-057-475B-4052
30	14	6.8	379	16	US-10-154-884B-4052
31	14	6.8	379	17	US-10-764-324-4052
32	14	6.8	394	9	US-09-960-352-11088
33	13	6.3	181	10	US-09-814-353-15407
34	13	6.3	425	10	US-09-814-353-2690
35	13	6.3	425	10	US-09-814-353-9023
36	13	6.3	439	10	US-09-825-751A-7
37	13	6.3	439	10	US-09-825-751A-33
38	13	6.3	468	9	US-09-960-352-10851
39	13	6.3	479	10	US-09-814-353-2217
40	13	6.3	479	10	US-09-814-353-8557
41	13	6.3	498	10	US-09-814-353-14941
42	13	6.3	546	17	US-10-283-975A-330
43	13	6.3	614	9	US-09-879-792-33
44	13	6.3	614	18	US-10-806-370-33
45	12	5.8	304	17	US-10-469-285-111

ALIGNMENTS

RESULT 1  
US-10-741-601-2633  
; Sequence 2633, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2633  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-2633

Score: 66.00 Matches: 66  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 31.88% Indels: 0  
 DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-2633 (1-201)

Qy 12 GluTrpValLeuThrAlaAlaHisCysLeuGluLysSerProArgProSerSerTyrllys 31  
 Db 3 GAGTGGGTGTTGACTGCTGCCACTGCTTGAGAGAGTCCCAAGGCCTTCATCCTCAAG 62

Qy 32 VallileLeuGlyAlaHisGlnGluValAsnLeuGluProHisValGlnGluLeuVal 51  
 Db 63 GTCATCTCTGGGTGCACACCAAGAGTGAATCTCGAACCCRCATGTTCCAGAAATAGAGTG 122

Qy 52 SerArgLeuPheLeuGluProThrArgLysAspIleAlaLeuLeuLysSerSerPro 71  
 Db 123 TCTAGGCTGTTCTTGGAGCCACACGAAAGATATTGCTTGTAAAGCTAAGCAGTCT 182

Qy 72 AlaVallileThrAspLys 77  
 Db 183 GCCGTCACTGACAAA 200

## RESULT 2

US-10-741-601-2649  
 ; Sequence 2649, Application US/10741601  
 ; Publication No. US20040166519A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2649  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-601-2649

Alignment Scores:  
 Pred. No.: 1.65e-60 Length: 201  
 Score: 66.00 Matches: 66  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 31.88% Indels: 0  
 DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-2649 (1-201)

Qy 127 AsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGluLeuCysAlaGlyHisLeu 146  
 Db 3 AATCGCTATGAGTTTCTGAATGGAAGAGTCCAAATCCACGAACTCTGCTGGGCAATTG 62

Qy 147 AlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLys 166  
 Db 63 GCCGAGGCACTGACAGTTGCCAGGTGACAGTGAGGKCTCTGTTGCTTCGAGAAG 122

Qy 167 AspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAsnLys 186  
 Db 123 GACAATAACATTTTACAGGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCAATAAG 182

Qy 187 ProGlyValTyrValArg 192  
 Db 183 CCTGGTGTCTATGTTCTG 200

## RESULT 3

US-10-741-601-2657  
 ; Sequence 2657, Application US/10741601  
 ; Publication No. US20040166519A1

; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2657  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-601-2657

Alignment Scores:  
 Pred. No.: 1.65e-60 Length: 201  
 Score: 66.00 Matches: 66  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 31.88% Indels: 0  
 DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-2657 (1-201)

Qy 59 ThrArgLysAspIleAlaLeuLeuLysSerSerProAlaVallileThrAspLysVal 78  
 Db 3 ACACGAAAGATATTGCTTGTAAAGCTAAGCAGTCTGCCGTCATCCTGACAAAGTA 62

Qy 79 IleProAlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIle 98  
 Db 63 ATCCAGCTTGTCTGCCATCCCCAAATATTGTGTCGTCGACCGACCGAATGTTTCATC 122

Qy 99 ThrGlyTrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeu 118  
 Db 123 ACTGGCTGGGAGAAACCCAAAGGTACTTTTGGAGTCTTTTGGAGTCTGCTTCTCAAGGAAGCCAGCTC 182

Qy 119 ProVallileGluAsnLys 124  
 Db 183 CCTGTGATTGAGAATAAA 200

## RESULT 4

US-10-741-601-2660  
 ; Sequence 2660, Application US/10741601  
 ; Publication No. US20040166519A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2660  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-601-2660

Alignment Scores:  
 Pred. No.: 1.65e-60 Length: 201  
 Score: 66.00 Matches: 66  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 31.88% Indels: 0  
 DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-2660 (1-201)

Qy 118 LeuProVallileGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGln 137  
 Db 3 CTCCTGTGATTGAGAAATAAGTGTGCAATCGTAATCGTATGAGTTTCTGAATGGAAGAGTCCAA 62

QY 138 SerThrGluLeuValCysPheGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSer 157  
 Db 63 TCACCGAAGTCTGTCTGGGCTTTGGCCGAGGACACGAGTTCAGGAGTGCAGT 122  
 QY 158 GlyGlyProLeuValCysPheGlyHisLeuAlaGlyThrAspSerCysGlnGlyValThrSerTrp 177  
 Db 123 GGAGGGCTCTGTCTGGTTCGAGAGGAGCAATACATTTACAGGAGTCACTTCTGG 182  
 QY 178 GlyLeuGlyCysAlaArg 183  
 Db 183 GGTCTTGGCTGTGCAGC 200

RESULT 5

US-10-741-601-2630  
 ; Sequence 2630, Application US/10741601  
 ; Publication No. US20040166519A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH STENOSIS AND USES THEREOF  
 ; FILE REFERENCE: CL001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2630  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-601-2630

Alignment Scores:  
 Pred. No.: 9,09e-47 Length: 201  
 Score: 53.00 Matches: 53  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 25.60% Indels: 0  
 DB: 17 Gaps: 0

US-09-992-095b-54 (1-207) x US-10-741-601-2630 (1-201)

QY 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyriLeuGlnGlyVal 174  
 Db 3 GGTGACAGTGGAGGGCTCTGGTTCTTCGAGAGGACAAATACATTTTACAGGAGTC 62  
 QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyriValArgValSer 194  
 Db 63 ACTTCTTGGGGTCTTGGCTGTGCAGCCCCCAATAAGCCVGGTGTCTATGTTCTGTTTCA 122  
 QY 195 ArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207  
 Db 123 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 161

RESULT 6

US-10-741-601-14280  
 ; Sequence 14280, Application US/10741601  
 ; Publication No. US20040166519A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH STENOSIS AND USES THEREOF  
 ; FILE REFERENCE: CL001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 14280  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-741-601-14280

Alignment Scores:

Pred. No.: 9,09e-47 Length: 201  
 Score: 53.00 Matches: 53  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 25.60% Indels: 0  
 DB: 17 Gaps: 0

US-09-992-095b-54 (1-207) x US-10-741-601-14280 (1-201)

QY 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyriLeuGlnGlyVal 174  
 Db 3 GGTGACAGTGGAGGGCTCTGGTTCTTCGAGAGGACAAATACATTTTACAGGAGTC 62  
 QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyriValArgValSer 194  
 Db 63 ACTTCTTGGGGTCTTGGCTGTGCAGCCCCCAATAAGCCVGGTGTCTATGTTCTGTTTCA 122  
 QY 195 ArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207  
 Db 123 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 161

RESULT 7

US-09-918-995-35150/c  
 ; Sequence 35150, Application US/09918995  
 ; Publication No. US20030073623A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED FROM VARIOUS CDNA LIBRARIES  
 ; FILE REFERENCE: 20411-756  
 ; CURRENT APPLICATION NUMBER: US/09/918,995  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: US/09/235,076  
 ; PRIOR FILING DATE: 1999-01-20  
 ; NUMBER OF SEQ ID NOS: 38054  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 35150  
 ; LENGTH: 508  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(508)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-918-995-35150

Alignment Scores:

Pred. No.: 2,11e-46 Length: 508  
 Score: 53.00 Matches: 53  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 25.60% Indels: 0  
 DB: 10 Gaps: 0

US-09-992-095b-54 (1-207) x US-09-918-995-35150 (1-508)

QY 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyriLeuGlnGlyVal 174  
 Db 427 GGTGACAGTGGAGGGCTCTGGTTCTTCGAGAGGACAAATACATTTTACAGGAGTC 368  
 QY 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyriValArgValSer 194  
 Db 367 ACTTCTTGGGGTCTTGGCTGTGCAGCCCCCAATAAGCCVGGTGTCTATGTTCTGTTTCA 308  
 QY 195 ArgPheValThrTrpIleGluGlyValMetArgAsnAsn 207  
 Db 307 AGGTTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 269

RESULT 8

US-10-741-601-14344  
 ; Sequence 14344, Application US/10741601  
 ; Publication No. US20040166519A1

```

; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14344
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-14344

Alignment Scores:
Pred. No.: 2,28e-39 Length: 201
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.22% Indels: 0
DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-14344 (1-201)

Qy 24 SerProArgProSerSerTyfLysValIleLeuGluValSerArgLeuPheLeuGluProThrArgLysAspIle 43
Db 39 TCCCCAGGCCCTTCATCCTACAGGTCTCCTGGGTGCACACCAAGAGTGAATCTCGAA 98

Qy 44 ProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArgLysAspIle 63
Db 99 CCRCATGTTCCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGAAAGATATT 158

Qy 64 AlaLeuLeuLysLeuSer 69
Db 159 GCCTTGCTAAAGCTAAGC 176

RESULT 9
US-10-741-601-14415
; Sequence 14415, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14415
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-14415

Alignment Scores:
Pred. No.: 6,49e-31 Length: 201
Score: 38.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.36% Indels: 0
DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-14415 (1-201)

Qy 155 GlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyfIleLeuGlnGlyVal 174
Db 87 GGTGACAGTGGAGGKCTCTGCTGTTTCTCGAAGAGGCAATATACATTTTACAGGATC 146

Qy 175 ThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyValArg 192
Db 147 ACTTCTTGGGGTCTTGCTGTGCACGCCCCCAATAAGCCCTGGTGTCTAATGTTCT 200

```

```

RESULT 10
US-10-741-601-14376
; Sequence 14376, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14376
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-14376

Alignment Scores:
Pred. No.: 7,4e-30 Length: 201
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.87% Indels: 0
DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-14376 (1-201)

Qy 118 LeuProValIleGluAenLysValCysAsnArgTyfGluPheLeuAenGlyArgValGln 137
Db 3 CTCCTGTGATTCAGATAAAGTGTGCATTCGATGATTTCTGAATGGAGAGTCAA 62

Qy 138 SerThrGluLeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGln 154
Db 63 TCCACCGAACTCTGTGCTGGCAITTTGCCGGAGGCACACAGTTGCCAG 113

RESULT 11
US-10-741-601-14492
; Sequence 14492, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741.601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14492
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-14492

Alignment Scores:
Pred. No.: 7,4e-30 Length: 201
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.87% Indels: 0
DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-14492 (1-201)

Qy 70 SerProAlaValIleThrAspLysValIleProAlaCysLeuProSerProAsnTyfVal 89
Db 36 AGTCCTCGCGTCATCAGTCAAAAGTAAATCCAGCTGTCTGCCATCCCAATATATGTG 95

Qy 90 ValAlaAspArgThrGluCysPheIleThrGlyTrpGlyGluThrGlnGly 106

```



Db 96 GTCGCGACCGGACCGAATGTTTCATCACTGGCTGGGAGAAACCAAGGT 146

RESULT 12

US-09-918-995-30128  
 ; Sequence 30128, Application US/09918995  
 ; Publication No. US20030073623A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 ; FILE REFERENCE: 20411-756  
 ; CURRENT APPLICATION NUMBER: US/09/918,995  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: US/09/235,076  
 ; PRIOR FILING DATE: 1999-01-20  
 ; NUMBER OF SEQ ID NOS: 38054  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 30128  
 ; LENGTH: 363  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)...(363)  
 ; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-30128

Alignment Scores:  
 Pred. No.: 1.44e-28 Length: 363  
 Score: 36.00 Matches: 36  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 17.39% Indels: 0  
 DB: 10 Gaps: 0

US-09-992-095b-54 (1-207) x US-09-918-995-30128 (1-363)

QY 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20  
 Db 255 ATGCACCTTCTGTGGAGGACCTTGATATCCCAAGAGTGGGTGTGACTGCTGCCACTGC 314  
 QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeuGlyAla 36  
 Db 315 TTGGAGAGTCCCAAGGCTTCATCTCAAGGTATCTCTGGGTGCC 362

RESULT 13

US-09-918-995-30172  
 ; Sequence 30172, Application US/09918995  
 ; Publication No. US20030073623A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 ; FILE REFERENCE: 20411-756  
 ; CURRENT APPLICATION NUMBER: US/09/918,995  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: US/09/235,076  
 ; PRIOR FILING DATE: 1999-01-20  
 ; NUMBER OF SEQ ID NOS: 38054  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 30172  
 ; LENGTH: 357  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens

US-09-918-995-30172

Alignment Scores:  
 Pred. No.: 1.05e-26 Length: 357  
 Score: 34.00 Matches: 34  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 16.43% Indels: 0  
 DB: 10 Gaps: 0

US-09-992-095b-54 (1-207) x US-09-918-995-30172 (1-357)

QY 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20  
 Db 255 ATGCACCTTCTGTGGAGGACCTTGATATCCCAAGAGTGGGTGTGACTGCTGCCACTGC 314  
 QY 21 LeuGluLysSerProArgProSerSerTyrLysValIleLeu 34  
 Db 315 TTGGAGAGTCCCAAGGCTTCATCTCAAGGTATCTCTGGGTGCC 356

RESULT 14

US-10-741-601-2639  
 ; Sequence 2639, Application US/10741601  
 ; Publication No. US20040166519A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE REFERENCE: CL001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2639  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens

US-10-741-601-2639

Alignment Scores:  
 Pred. No.: 1.25e-25 Length: 201  
 Score: 33.00 Matches: 65  
 Percent Similarity: 97.01% Conservative: 0  
 Best Local Similarity: 97.01% Mismatches: 1  
 Query Match: 15.94% Indels: 2  
 DB: 17 Gaps: 0

US-09-992-095b-54 (1-207) x US-10-741-601-2639 (1-201)

QY 59 ThrArgLysAspIleAlaLeuLysLeuSerProAlaValIleThrAspLysVal 78  
 Db 2 ACACGAAAAGATATTGCTTGTCTAAAGCTAAGAGTCTGCGCTCATCTGCAAAAGTA 61  
 QY 79 IleProAlaCysLeuProSerProAsnTyrValAlaAsp-ArgThrGluCysPheI 98  
 Db 62 ATCCAGCTTGTCTGCCATCCCAATATATGTGCTGCTRA-CCGAGCCGAATGTTTCAT 120  
 QY 98 eThrGlyTrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLe 118  
 Db 121 CACTGGCTGGGAGAAACCAAGGTACTTTTGGAGCTGGCCTTCTCAAGGAGCCAGCT 180

RESULT 15

US-10-741-601-2653  
 ; Sequence 2653, Application US/10741601  
 ; Publication No. US20040166519A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE REFERENCE: CL001500  
 ; CURRENT APPLICATION NUMBER: US/10/741,601  
 ; CURRENT FILING DATE: 2003-12-22  
 ; NUMBER OF SEQ ID NOS: 26415  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2653  
 ; LENGTH: 201  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens

US-10-741-601-2653

Alignment Scores:  
 Pred. No.: 1.25e-25 Length: 201  
 Score: 33.00 Matches: 65  
 Percent Similarity: 97.01% Conservative: 0  
 Best Local Similarity: 97.01% Mismatches: 1  
 Query Match: 15.94% Indels: 2  
 DB: 17 Gaps: 0

US-09-992-095B-54 (1-207) x US-10-741-601-2653 (1-201)

Qy	58	ProThrArgLysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLys	77
Db	2	CCCAACACGAAAGATATTGGCTTGCTAAAGCTAAGCAGTCCTGCCGTCATCACTGACAAA	61
Qy	78	ValIleProAlaCysLeuProSerProAsnTyrValValAla-AspArgThrGluCysPh	97
Db	62	GTAATCCCAGCTTGCTGCCATCCCAATTTATGTGTCRC-TGACCGGACCGAATGTTT	120
Qy	97	eileThrGlyTrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLysGluAlaGl	117
Db	121	CATCACTGGCTGGGGAGAAACCAAGGTACTTTTGGAGCTGGCTTCTCAAGGAAGCCCA	180
Qy	117	nLeuProValIleGluAsn	123
Db	181	GCTCCCTGTGATTGAGAAT	199

Search completed: October 27, 2004, 12:58:47  
 Job time : 496 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:20:12 ; Search time 32 Seconds  
(without alignments)  
622.402 Million cell updates/sec

Title: US-09-992-095B-54  
Perfect score: 207  
Sequence: 1 MHFCGGTLISPEWVLTAAHC.....GVYVRVSFRVTWIEGVMRNN 207

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105624

Minimum DB seq length: 0  
Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Database :

PIR 79: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	5.8	126	A23473	chymotrypsin-like
2	10	4.8	30	A32946	trypsin-like serin
3	10	4.8	61	PS0049	serine proteinase
4	10	4.8	149	S35208	serine proteinase
5	9	4.3	73	S44462	elastase (EC 3.4.2
6	9	4.3	75	A37002	catroboxin I (EC 3
7	9	4.3	191	S54115	complement factor
8	9	4.3	196	T08808	hypothetical prote
9	8	3.9	30	C32946	serine proteinase
10	8	3.9	43	A61168	cocoonase (EC 3.4.
11	8	3.9	81	A18966	tissue kallikrein
12	8	3.9	85	S44461	elastase (EC 3.4.2
13	8	3.9	90	JE0210	proteinase (EC 3.4
14	8	3.9	94	PC2013	tissue kallikrein
15	8	3.9	161	I62744	coagulation factor
16	8	3.9	161	I48158	coagulation factor
17	7	3.4	25	A24807	cytotoxic T-lympho
18	7	3.4	30	A61333	trypsin (EC 3.4.21
19	7	3.4	40	A49081	capillary permeabi
20	7	3.4	66	I52972	kallikrein - mouse
21	7	3.4	96	A05308	tissue kallikrein
22	7	3.4	97	AF1451	transcription regu
23	7	3.4	104	S15395	tissue kallikrein-
24	7	3.4	137	S55364	serine proteinase
25	7	3.4	149	1 QOMSM	tissue kallikrein
26	7	3.4	151	2 S35205	proteinase 5 - buf
27	7	3.4	152	2 S35209	serine proteinase
28	7	3.4	152	2 S35206	serine proteinase
29	7	3.4	152	2 S35203	serine proteinase

30	7	3.4	156	2 B23863	tissue kallikrein
31	7	3.4	158	2 S35201	serine proteinase
32	7	3.4	171	2 B69345	hypothetical prote
33	7	3.4	177	2 S23505	chymase (EC 3.4.21
34	7	3.4	188	2 B32340	tissue kallikrein
35	7	3.4	190	2 D75006	hypothetical prote
36	6	2.9	22	B32946	serine proteinase
37	6	2.9	30	E32946	serine proteinase
38	6	2.9	30	D32946	serine proteinase
39	6	2.9	38	A32790	interferon gamma-1
40	6	2.9	67	1 C37873	allopheycocyanin li
41	6	2.9	73	2 D50219	LSU ribosomal prot
42	6	2.9	73	2 H72408	hypothetical prote
43	6	2.9	77	2 H86679	prophage pil prote
44	6	2.9	80	2 A05324	gamma-tenin (EC 3.
45	6	2.9	80	2 A64382	ribosomal protein

ALIGNMENTS

RESULT 1

A3473  
chymotrypsin-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)  
N:Alternate names: pancreatic elastase II [misidentification]  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004

C:Accession: A23473  
R:Vered, M.; Gertler, A.; Burstein, Y.  
Int. J. Pept. Protein Res. 27, 183-190, 1986  
A:Reference number: A23473; MUID:86194934; PMID:3634756  
A:Accession: A23473  
A:Molecule type: protein  
A:Residues: 1-126 <VER>  
A:Cross-references: UNIPROT:Q7M325  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase

Query Match 5.8%; Score 12; DB 2; Length 126;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SCQGDGGGLVLC 163  
DB 95 SCQGDGGGLVLC 106  
|||||

RESULT 2

A32946  
trypsin-like serine proteinase (EC 3.4.21.-) 1 - nematode (Anisakis simplex) (fragments)  
C:Species: Anisakis simplex  
C:Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 20-Sep-1999  
C:Accession: A32946  
R:Sakanari, J.A.; Staunton, C.E.; Eakin, A.E.; Craik, C.S.; McKerrow, J.H.  
Proc. Natl. Acad. Sci. U.S.A. 86, 4863-4867, 1989  
A:Title: Serine proteases from nematode and protozoan parasites: isolation of sequence  
A:Reference number: A32946; MUID:89296904; PMID:2662185

A:Accession: A32946  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr  
A:Molecule type: DNA  
A:Residues: 1-30 <SAK>  
C:Keywords: hydrolase; serine proteinase

Query Match 4.8%; Score 10; DB 2; Length 30;  
Best Local Similarity 100.0%; Pred. No. 0.012; 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 151 DSCQGDGGGP 160  
DB 21 DSCQGDGGGP 30  
|||||

RESULT 3

PS0049  
serine proteinase (EC 3.4.21.-) 3 - fruit fly (*Drosophila melanogaster*) (fragment)  
C:Species: *Drosophila melanogaster*  
C:Date: 07-Jun-1990 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C:Accession: PS0049  
R:Yun, Y.; Davis, R.L.  
Mol. Cell. Biol. 9, 692-700, 1989  
A:Title: Levels of RNA from a family of putative serine protease genes are reduced in *Drosophila* larvae  
A:Reference number: JS0260; MUID:89219063; PMID:2469005  
A:Accession: PS0049  
A:Molecule type: DNA  
A:Residues: 1-61 <YUN>  
A:Cross-references: UNIPROT:P17207; GB:M24380; NID:g158412  
A:Experimental source: strain dnc mutant  
C:Genetics:  
A:Gene: SER3  
A:Cross-references: FlyBase:FBgn0003358  
A:Map position: 3 99C-D  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-53/Domain: trypsin homology (fragment) <TRY>  
F:30/Binding site: substrate (Val) #status Predicted

Query Match 4.8%; Score 10; DB 2; Length 61;  
Best Local Similarity 100.0%; Pred. No. 0.022; Mismatches 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0;

Qy 153 CQGDGGGGLV 162  
Db 7 CQGDGGGGLV 16  
|||||

RESULT 4  
S35208  
serine proteinase (EC 3.4.21.-) 8 - buffalo fly (fragment)  
C:Species: *Haematobia irritans exigua* (buffalo fly)  
C:Date: 10-Dec-1993 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C:Accession: S35208; S42696  
R:Elvin, C.M.; Whan, V.; Riddles, P.W.  
Mol. Gen. Genet. 240, 132-139, 1993  
A:Title: A family of serine protease genes expressed in adult buffalo fly (*Haematobia irritans*)  
A:Reference number: S35201; MUID:93341451; PMID:8341258  
A:Accession: S35208  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-149 <ELV>  
A:Cross-references: UNIPROT:Q06784; EMBL:222567  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-149/Domain: trypsin homology (fragment) <TRY>  
F:7,52,146/Active site: His, Asp, Ser #status Predicted

Query Match 4.8%; Score 10; DB 2; Length 149;  
Best Local Similarity 100.0%; Pred. No. 0.046; Mismatches 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0;

Qy 151 DSCQDGGG 160  
Db 140 DSCQDGGG 149  
|||||

RESULT 5  
S44462  
elastase (EC 3.4.21.-) 2B - horse (fragments)  
C:Species: *Equus caballus* (domestic horse)  
C:Date: 20-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 09-Jul-2004  
C:Accession: S44462  
R:Dubin, A.; Potempa, J.; Travis, J.  
Biochem. J. 300, 401-406, 1994  
A:Title: Structural and functional characterization of elastases from horse neutrophils.  
A:Reference number: S44461; MUID:94271153; PMID:7516152  
A:Accession: S44462  
A:Molecule type: protein

A:Residues: 1-31;32-56;57-73 <DUB>  
A:Cross-references: UNIPROT:P37358  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-73/Domain: trypsin homology (fragments) <TRY>

Query Match 4.3%; Score 9; DB 2; Length 73;  
Best Local Similarity 100.0%; Pred. No. 0.24; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 155 GDSGGPLVC 163  
Db 62 GDSGGPLVC 70  
|||||

## RESULT 6

A37002  
catroxin I (EC 3.4.21.-) - western diamondback rattlesnake (fragments)  
C:Species: *Crotalus atrox* (western diamondback rattlesnake)  
C:Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 09-Jul-2004  
C:Accession: A37002  
R:Pirkle, H.; Theodor, I.; Lopez, R.  
Thromb. Res. 56, 159-168, 1989  
A:Title: Catroxin, a weakly thrombin-like enzyme from the venom of *Crotalus atrox*. NH<sub>2</sub>-terminal amino acid sequence and primary structure.  
A:Reference number: A37002; MUID:90141479; PMID:2617466  
A:Accession: A37002  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-75 <PIR>  
A:Cross-references: UNIPROT:Q7LZF5  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase; venom  
F:1-75/Domain: trypsin homology (fragments) <TRY>

Query Match 4.3%; Score 9; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 0.25; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 12 EWLTAHHC 20  
Db 34 EWLTAHHC 42  
|||||

## RESULT 7

S54115  
complement factor D (EC 3.4.21.46) - pig (fragment)  
C:Species: *Sus scrofa domestica* (domestic pig)  
C:Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: S54115  
R:Nicolas, N.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: S54115  
A:Accession: S54115  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-191 <NIC>  
A:Cross-references: UNIPROT:P51779; EMBL:Z49058; NID:g773264; PIDN:CAA88844.1; PID:g773264  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase  
F:1-181/Domain: trypsin homology (fragment) <TRY>

Query Match 4.3%; Score 9; DB 2; Length 191;  
Best Local Similarity 100.0%; Pred. No. 0.53; Mismatches 0; Indels 0; Gaps 0;  
Matches 9; Conservative 0;

Qy 155 GDSGGPLVC 163  
Db 139 GDSGGPLVC 147  
|||||

## RESULT 8

T08808  
hypothetical protein DKFZp586J1923.1 - human (fragment)

C:Species: Homo sapiens (man)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: T08808  
 R:Ansgorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, May 1999  
 A:Reference number: Z16472  
 A:Accession: T08808  
 A:Molecule type: mRNA  
 A:Residues: 1-196 <ANS>  
 A:Cross-references: UNIPROT:Q9UKR3; EMBL:AL050220  
 A:Experimental source: adult uterus; clone DKFZp586L1923  
 C:Genetics:  
 A:Note: DKFZp586J1923.1  
 C:Superfamily: trypsin; trypsin homology

Query Match 4.3%; Score 9; DB 2; Length 196;  
 Best Local Similarity 100.0%; Pred. No. 0.54;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 WVLTAHCL 21  
 |||||  
 DB 86 WVLTAHCL 94

RESULT 9  
 C32946  
 serine proteinase (EC 3.4.21.-) 3 - nematode (Anisakis simplex) (fragments)  
 C:Species: Anisakis simplex  
 C:Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 20-Sep-1999  
 C:Accession: C32946  
 R:Sakanari, J.A.; Staunton, C.E.; Eakin, A.E.; Craik, C.S.; McKerrow, J.H.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 4863-4867, 1989  
 A:Title: Serine proteases from nematode and protozoan parasites: isolation of sequence H  
 A:Reference number: A32946; MUID:89296904; PMID:2662185  
 A:Accession: C32946  
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
 A:Molecule type: DNA  
 A:Residues: 1-30 <SAK>  
 C:Keywords: hydrolase; serine proteinase

Query Match 3.9%; Score 8; DB 2; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 CQDGGGP 160  
 |||||  
 DB 23 CQDGGGP 30

RESULT 10  
 A61168  
 cocoanase (EC 3.4.21.-) precursor - polyphemus moth (fragments)  
 C:Species: Anthraea polyphemus (polyphemus moth)  
 C:Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 09-Jul-2004  
 C:Accession: A61168  
 R:Kramer, K.O.; Felsted, R.L.; Law, J.H.  
 J. Biol. Chem. 248, 3021-3028, 1973  
 A:Title: Cocoanase. V. Structural studies on an insect serine protease.  
 A:Reference number: A61168; MUID:73166540; PMID:4735570  
 A:Accession: A61168  
 A:Molecule type: protein  
 A:Residues: 1-43 <KRA>  
 A:Cross-references: UNIPROT:Q7M3M4  
 C:Keywords: hydrolase; serine proteinase; zymogen  
 F:1-13/Domain: activation peptide #status predicted <APT>  
 F:14-43/Product: cocoanase (fragment) #status experimental <MAT>

Query Match 3.9%; Score 8; DB 2; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 CQDGGGP 160  
 |||||

DB 30 CQDGGGP 37

# RESULT 11

A18966  
 tissue kallikrein (EC 3.4.21.35) mCK-22 - mouse (fragments)  
 N:Alternate names: epidermal growth factor-binding protein type A; nerve growth factor b  
 C:Species: Mus musculus (house mouse)  
 C:Date: 31-Mar-1998 #sequence\_revision 06-Feb-1995 #text\_change 18-Jul-1997  
 C:Accession: A18966; A38356; C18966  
 R:Anundi, H.; Ronne, H.; Peterson, P.A.; Rask, L.  
 Eur. J. Biochem. 129, 365-371, 1982  
 A:Title: Partial amino-acid sequence of the epidermal growth-factor-binding protein.  
 A:Reference number: A91126; MUID:83105150; PMID:6295764  
 A:Accession: A18966  
 A:Molecule type: protein  
 A:Residues: 1-81 <ANU>  
 R:Fannestock, M.; Woo, J.E.; Lopez, G.A.; Snow, J.; Walz, D.A.; Arici, M.J.; Mobley, W.  
 Biochemistry 30, 3443-3450, 1991  
 A:Title: Beta-NGF-endorpeptidase: structure and activity of a kallikrein encoded by the  
 A:Reference number: A38356; MUID:91190897; PMID:2012805  
 A:Accession: A38356  
 A:Molecule type: protein  
 A:Residues: 1-30 <FAH>  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:1-81/Domain: trypsin homology (fragments) <TRY>

Query Match 3.9%; Score 8; DB 2; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 2.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 WVLTAHCL 20  
 |||||  
 DB 35 WVLTAHCL 42

# RESULT 12

S44461  
 elastase (EC 3.4.21.-) 2A - horse (fragments)  
 C:Species: Equus caballus (domestic horse)  
 C:Date: 20-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 09-Jul-2004  
 C:Accession: S44461  
 R:Dubin, A.; Potempa, J.; Travis, J.  
 Biochem. J. 300, 401-406, 1994  
 A:Title: Structural and functional characterization of elastases from horse neutrophils  
 A:Reference number: S44461; MUID:94271153; PMID:7516152  
 A:Accession: S44461  
 A:Molecule type: protein  
 A:Residues: 1-34;35-59;60-85 <DUB>  
 A:Cross-references: UNIPROT:P37357  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase

Query Match 3.9%; Score 8; DB 2; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 2.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 HFCGGTLLI 9  
 |||||  
 DB 27 HFCGGTLLI 34

# RESULT 13

JE0210  
 proteinase (EC 3.4.-.-) serine-like, NBS1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 07-May-1999  
 C:Accession: JE0210  
 R:Juo, L.; Herbrick, J.A.; Scherer, S.W.; Beatty, B.; Squire, J.; Diamandis, E.P.  
 Biochem. Biophys. Res. Commun. 247, 580-586, 1998  
 A:Title: Structure characterization and mapping of the normal epithelial cell-specific  
 A:Reference number: JE0210; MUID:98321170; PMID:9647736

QY 153 CQGDSSGP 160  
Db 119 CQGDSSGP 126

Search completed: October 27, 2004, 09:32:35  
Job time : 34 secs

A;Accession: JE0210  
A;Molecule type: mRNA  
A;Residues: 1-90 <LUO>  
A;Cross-references: GB:AF055481  
C;Genetics:  
A;Gene: GDB:PRSSL1; NES1  
A;Map position: 19q13.3-13.4  
C;Keywords: hydrolase

Query Match 3.9%; Score 8; DB 2; Length 90;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 DSGGPLVC 163  
Db 74 DSGGPLVC 81

RESULT 14  
PC2013  
tissue kallikrein (EC 3.4.21.35) mK1, submandibular - mouse (fragments)  
N;Alternate names: proteinase F  
C;Species: Mus musculus (house mouse)  
C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 09-Jul-2004  
C;Accession: PC2013  
R;Hosoi, K.; Tsunashawa, S.; Kurihara, K.; Aoyama, H.; Ueha, T.; Murai, T.; Sakiyama, F.  
J. Biochem. 115, 137-143, 1994  
A;Title: Identification of mK1, a true tissue (glandular) kallikrein of mouse submandibular gland  
A;Reference number: PC2013; MUID:94245648; PMID:8188620  
A;Accession: PC2013  
A;Molecule type: protein  
A;Residues: 1-94 <HOS>  
A;Cross-references: UNIPROT:Q7MOB4  
A;Experimental source: submandibular gland  
C;Genetics:  
A;Gene: mK1  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase; submandibular gland  
F;1-94/Domain: trypsin homology (fragments) <TRY>

Query Match 3.9%; Score 8; DB 2; Length 94;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 WVLTAHC 20  
Db. 35 WVLTAHC 42

RESULT 15  
I62744  
coagulation factor Xa (EC 3.4.21.6) - rhesus macaque (fragment)  
C;Species: Macaca mulatta (rhesus macaque)  
C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 09-Jul-2004  
C;Accession: I62744  
R;Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.  
Eur. J. Haematol. 52, 162-168, 1994  
A;Title: Analysis of the partial nucleotide sequences and deduced primary structures of  
A;Reference number: I46196; MUID:94222160; PMID:8168596  
A;Accession: I62744  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-161 <RES>  
A;Cross-references: UNIPROT:Q28511; GB:D21214; NID:9415307; PIDN:BAA04755.1; PID:9455395  
C;Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
C;Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxylglutamate  
F;1-161/Domain: trypsin homology (fragment) <TRY>

Query Match 3.9%; Score 8; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 27, 2004, 09:38:24 ; Search time 3062 Seconds  
(without alignments)  
2463.431 Million cell updates/sec

Title: US-09-992-095b-54  
Perfect score: 207  
Sequence: 1 MHFCGGLISPEWVLTAAHC.....GVYVRVSRFTWIEGVMRNN 207

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues  
Word size: 1  
Total number of hits satisfying chosen parameters: 41268600

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cpn2.1/USPTO.spool/US09992095/runat.26102004.090253.11815/app.query.fasta\_1.391  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=621  
-USER=US09992095 @CGN 1.1 3437 @runat.26102004.090253.11815 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hic.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_gss1.\*  
9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	73.4	523	6	CB164684
2	116	56.0	616	5	BU073846
3	114	55.1	433	7	H90220
4	113	54.6	611	1	AV662084
5	106	51.2	564	6	CB161593
6	105	50.7	475	7	N91337
7	104	50.2	517	6	CB162370
8	104	50.2	543	6	CB161910
9	103	49.8	481	7	H60805

100	48.3	594	1	AV661991	AV661991
98	47.3	381	7	H37877	H37877
96	46.4	575	7	T84554	T84554
95	45.9	504	7	N77239	N77239
94	45.9	541	1	AA722885	AA722885
92	44.4	593	2	BE348267	BE348267
90	43.5	595	2	AW950594	AW950594
89	43.0	459	7	R94305	R94305
88	42.5	563	2	AW271976	AW271976
87	42.0	303	7	T73187	T73187
86	42.0	424	7	H73861	H73861
85	40.1	614	1	AV693554	AV693554
84	39.1	621	1	AI377474	AI377474
83	38.6	405	2	AW950595	AW950595
82	38.2	464	2	BE326689	BE326689
81	36.7	368	7	H95494	H95494
80	36.7	457	1	AI631756	AI631756
79	36.7	466	1	AI948806	AI948806
78	36.7	558	2	AW104579	AW104579
77	36.7	545	1	AV720081	AV720081
76	32.9	341	7	T70507	T70507
75	31.9	462	1	AI768449	AI768449
74	31.4	553	1	AI633979	AI633979
73	30.4	444	1	AI640315	AI640315
72	30.4	589	7	H73871	H73871
71	29.5	444	7	T73867	T73867
70	29.5	462	1	AA995929	AA995929
69	29.0	360	1	AV654539	AV654539
68	29.0	607	1	AV662061	AV662061
67	27.5	393	7	T51771	T51771
66	26.1	416	7	T68303	T68303
65	26.1	515	1	AI351543	AI351543
64	25.6	423	4	BG565038	BG565038
63	25.6	435	2	BF362261	BF362261
62	25.6	549	2	BF090010	BF090010
61	25.6	562	2	BF952026	BF952026

ALIGNMENTS

RESULT 1  
LOCUS CB164684 523 bp mRNA linear EST 30-JAN-2003  
DEFINITION K-EST0225947 L17N670205n1 Homo sapiens cDNA clone  
CB164684 L17N670205n1-43-C07 5', mRNA sequence.  
CB164684.1 GI:28150810  
EST.  
CB164684.1 Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 523)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 43 row: C column: 07  
High quality sequence stop: 523.

FEATURES  
source  
1..523  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L17N670205n1-43-C07"

```

/sex="F"
/lab host="Top10F"
/clone lib="L17N670205n1"
/notes="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
Site 2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN
Alignment Scores:
Pred. No.: 3,77e-140 Length: 523
Score: 152.00 Matches: 152
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.43% Indels: 0
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CB164684 (1-523)
QY 56 LeuGluProThrArgLysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThr 75
DB 30 CTGGAGCCACACGAAAGATATTGCCTTGCTAAGCTAAGCAGTCTCTGCGTCATCACT 89
QY 76 AspLysValIleProAlaCysLeuProSerProLeuTyrValValAlaAspArgThrGlu 95
DB 90 GACAAAGTAATCCCGAGTGTGTGCGATCCCAAAATATATGTGTGCTGACCGGACCGAA 149
QY 96 CysPheIleThrGlyTrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGlu 115
DB 150 TGTTCATCACTGCGTGGGGAGAACCCCAAGTACTTTTGGAGCTGGGCTTCTCAAGGAA 209
QY 116 AlaGlnLeuProValIleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArg 135
DB 210 GCCAGCTCCCTGTGATTGAGAAATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGA 269
QY 136 ValGlnSerThrGluLeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGly 155
DB 270 GTCCAAATCCAGCAACTCTGTGCTGGGCAATTTGGCCGGAGGCATGACAGTTGCCAGGGT 329
QY 156 AspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThr 175
DB 330 GACATGTGAGGGCCCTCTGTTTGTCTCGAGAGGACAAATACATTTTACAGGAGTCACT 389
QY 176 SerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyrValArgValSerArg 195
DB 390 TCTTGGGGTCTGGCTGTGCACGCCCAATAAGCTGTGTCTATGTTCGTGTTTCAAGG 449
QY 196 PheValThrTrpIleGluGlyValMetArgAsn 207
DB 450 TTTGTTACTTGGATTGAGGGAGTGATGAGAAATAAT 485

RESULT 2
BU073846
LOCUS in25c11.y1 Human Fetal Pancreas 1B Homo sapiens cDNA clone IMAGE:
DEFINITION 5', similar to SW:PLMN_HUMAN P00747 PLASMINOGEN PRECURSOR ;, mRNA
sequence.
ACCESSION BU073846
VERSION BU073846.1 GI:22515035
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 616)
AUTHORS Melkonian, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Teagareishvili, R.,
Williams, T., Jackson, Y. and Bowers, Y.

TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohpc.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@imgate.wustl.edu)
Putative full length read
vector to vector length is 617
Seq primer: -40RP from Gibco
High quality sequence stop: 470.

FEATURES
Location/Qualifiers
1..616
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:"
/tissue_type="Fetal Pancreas (4 Pooled Donors, 18 - 20
weeks, Stratagene #738023)"
/dev_stage="Fetal Pancreas"
/clone_lib="Human Fetal Pancreas 1B"
/notes="vector: pBluescript SK(-); Site 1: NotI; Site 2:
XhoI; cDNA made by oligo-dT priming. Size-selected on
agarose gel. Average insert size ~1kb. 5' XhoI site was
destroyed after directional cloning. Amplified once.
Contact information: Hiroshi Inoue, MD, Metabolism Div.
(Alan Permutt Lab), Washington University School of
Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO
63110 E-mail: hinoue@imgate.wustl.edu Tel: 314-362-1916,
Fax:314-747-3692."

ORIGIN
Alignment Scores:
Pred. No.: 2,09e-104 Length: 616
Score: 116.00 Matches: 116
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 56.04% Indels: 0
DB: 5 Gaps: 0

US-09-992-095B-54 (1-207) x BU073846 (1-616)
QY 92 AspArgThrGluCysPheIleThrGlyTrpGlyGluThrGlnGlyThrPheGlyValaGly 111
DB 3 GACCGGACCGGAATGTTTTCATCCTGCTGCTGGGGAGAAACCCCAAGTACTTTTGGAGCTGCG 62
QY 112 LeuLeuLysGluAlaGlnLeuProValIleGluAsnLysValCysAsnArgTyrGluPhe 131
DB 63 CTTTCTCAAGGAGGCCAGCTCCCTGTGATTGAGAAATAAAGTGTGCAATCGCTATGAGTTT 122
QY 132 LeuAsnGlyArgValGlnSerThrGluLeuCysAlaGlyHisLeuAlaGlyGlyThrAsp 151
DB 123 CTGAATGAGAGAGTCCAATCCACCGAATCTGTGCTGGGCATTTGGCCGGAGGCACTGAC 182
QY 152 SerCysGlnGlyAspSerGlyProLeuValCysPheGluLysAspLysTyrIleLeu 171
DB 183 AGTTGCCAGGGTGACAGTGGAGGTCCTCTGGTTGCTTCGAGAGGACAAATACATTTTA 242
QY 172 GlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyrVal 191
DB 243 CAAGGAGTCACTCTTGGGGTCTTGGCTGTGCACGCCCAATAAGCTGTGTCTATGTT 302
QY 192 ArgValSerArgPheValThrTrpIleGluGlyValMetArgAsn 207
DB 303 CGTGTTCAGAGTTTGTGTACTTGGATTGAGGGAGTGATGAGAAATAAT 350

```



**RESULT 3**

H90220 433 bp mRNA linear EST 28-NOV-1995  
LOCUS yu85g08.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
DEFINITION IMAGE:240638 5' similar to gb:X05199 PLASMINOGEN PRECURSOR  
(HUMAN); mRNA sequence.

ACCESSION H90220 GI:1080650  
VERSION H90220.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marzta,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
Insert Size: 1021  
High quality sequence stops: 355  
Source: IMAGE Consortium, LNLN  
This clone is available royalty-free through LNLN; contact the  
IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
Insert Length: 1021 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 355.

FEATURES  
source  
1..433  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3789771"  
/db\_xref="taxon:9606"  
/clone="IMAGE:240638"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/notes="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5'-AACTGGGAAGTAATTAAGAATCTTTTTCCTTTTTTTT 3'] double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.46e-102 Length: 433  
Score: 114.00 Matches: 114  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 55.07% Indels: 0  
DB: 7 Gaps: 0

US-09-992-095B-54 (1-207) x H90220 (1-433)

Qy 7 ThrLeuSerProGluTrpValLeuThrAlaHisCysLeuGluYsSerProArg 26  
Db 1 ACCTTGATATCCCAAGTAGTGGGTGTTGACTGTGCCACTGCTTGGAGAGTCCCCAAG 60  
27 ProSerSerTyrlsValIleLeuGlyAlaHisGlnGluValAsnLeuGluProHisVal 46

61 CCTTCATCTCAAGAGTGCATCTCGGTGCACACCAAGAGTGAATCTCGAACCCCATGTT 121

Qy 47 GlnGlulleGluValSerArgLeuPheLeuGluProThrArgLysAspIleAlaLeuLeu 66  
Db 121 CAGGAATAAGAGTGTCTAGGCTGTCTTGGAGCCCACACGAAAAGATATTGCCCTTGCTA 180

Qy 67 LysLeuSerSerProAlaValIleThrAspLysValIleProAlaCysLeuProSerPro 86  
Db 181 AGCTAAGCAGTCTCTGCCGTCACTGACAAGTAATCCAGCTTGTCTGCCATCCCA 240

Qy 87 AsnTyrrValValalaAspArgThrGluCysPheIleThrGlyTrpGlyGluThrGlnGly 106  
Db 241 AATTATGTGTGTCTGCTGACCGGACCAATGTTTTCATCACTGGCTGGGAGAAAACCAAGT 300

107 ThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120  
301 ACTTTTGGAGCTGGCCTTCTCAAGGAAGCCAGCTCCCTGTG 342

RESULT 4  
AV662084 611 bp mRNA linear EST 16-JAN-2000  
LOCUS AV662084 GLC Homo sapiens cDNA clone GLCHAG06 3', mRNA sequence.  
DEFINITION AV662084  
ACCESSION AV662084  
VERSION AV662084.1 GI:9883098  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X., Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W., Shen,K., Lu,G., Fu,S., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X., Hu,G., Gu,J., Chen,Z. and Han,Z.

TITLE Insight into hepatocellular carcinogenesis at transcriptome level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
MEDLINE 21625106  
PubMed 11752456

COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919 (ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers  
1..611  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GLCHAG06"  
/tissue\_type="corresponding non cancerous liver tissue"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_lib="GLC"  
/notes="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.95e-101 Length: 611  
Score: 113.00 Matches: 113  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 54.59% Indels: 0  
DB: 1 Gaps: 0

US-09-992-095B-54 (1-207) x AV662084 (1-611)

Qy 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20

```

Db      69 ATGCACCTTCTGTGGAGGACCTTATATCCCAAGATGGGTGTGACTGTGCCACTGC 128
Qy      21 LeuGluYsSerProArgProSerSerTyrIysValIleLeuGlyAlaHisGlnGluVal 40
Db      129 TTGGAGAAGTCCCAAGGCTTCATCTACAAAGTTCATCTGGGTGCACACCAAGAAGTG 188
Qy      41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db      189 AATCTCGAACCGCATGTTTCAGAAATAGAAGTGTCTAGGCTGTCTTGGAGCCACACGA 248
Qy      61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db      249 AAGATATTGCTTGTCTAAAGCTAAGCAGTCTCCGTCATCATGACAAAGTAATCCCA 308
Qy      81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db      309 GCTTGTCTGCCATCCCAAAATATGTTGGTCCGACCGGACCGAATGTTTCATCACTGGC 368
Qy      101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeu 113
Db      369 TGGGAGAAACCAAGGTACTTTTGGAGCTGGCTCTTA 407

```

```

RESULT 5
LOCUS      CB161593              564 bp      mRNA      linear      EST 30-JAN-2003
DEFINITION K-EST0221565 L17N670205n1 Homo sapiens cDNA clone
VERSION     L17N670205n1-35-F05 5', mRNA sequence.
ACCESSION   CB161593
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 564)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,Y.E., Sohn,H.Y., Kim,J.N., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.krribb.re.kr
            Plate: 35 row: F column: 05
            High quality sequence stop: 564.
            Location/Qualifiers
                1..564
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="L17N670205n1-35-F05"
                /sex="F"
                /lab_host="Top10F"
                /clone_lib="L17N670205n1"
                /note="Organ: Liver; Vector: pT7T3-Pac; Site 1: ECoRI;
                Site 2: NotI; The library was contributed by the Soares
                laboratory and it was constructed as described by Bonaldo,
                M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                6 (9): 791-806. RNA was prepared from harvested cell
                culture."

```

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.krribb.re.kr

Plate: 35 row: F column: 05

High quality sequence stop: 564.

Location/Qualifiers

1..564

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="L17N670205n1-35-F05"

/sex="F"

/lab\_host="Top10F"

/clone\_lib="L17N670205n1"

/note="Organ: Liver; Vector: pT7T3-Pac; Site 1: ECoRI;

Site 2: NotI; The library was contributed by the Soares

laboratory and it was constructed as described by Bonaldo,

M.F., Lennon, G. and Soares, M.B. (1996), Genome Research

6 (9): 791-806. RNA was prepared from harvested cell

culture."

# ORIGIN

```

Alignment Scores:
Pred. No.:      1..58e-94      Length:      564
Score:          106.00      Matches:      149
Percent Similarity: 98.03%      Conservative: 0
Best Local Similarity: 98.03%      Mismatches: 2
Query Match:    51.21%      Indels:      3

```

```

DB:      6      Gaps:      0
US-09-992-095B-54 (1-207) x CB161593 (1-564)
Qy      1 MetHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
Db      111 ATGCACCTTCTGTGGAGGACCTTATATCCCAAGATGGGTGTGACTGTGCCACTGC 170
Qy      21 LeuGluYsSerProArgProSerSerTyrIysValIleLeuGlyAlaHisGlnGluVal 40
Db      171 TTGGAGAAGTCCCAAGGCTTCATCTACAAAGTTCATCTGGGTGCACACCAAGAAGTG 230
Qy      41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db      231 AATCTCGAACCGCATGTTTCAGAAATAGAAGTGTCTAGGCTGTCTTGGAGCCACACGA 290
Qy      61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db      291 AAGATATTGCTTGTCTAAAGCTAAGCAGTCTCCGTCATCATGACAAAGTAATCCCA 350
Qy      81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db      351 GCTTGTCTGCCATCCCAAAATATGTTGGTCCGACCGGACCGAATGTTTCATCACTGGC 410
Qy      101 TrpGlyGluThrGlnGlyThrPhe-GlyAlaGlyLeuLeuLysGluAlaGlnLeuProVa 120
Db      411 TGGGAGAAACCAAGGTCT-TTTGGAGCTGGCCTTCTCAAGGAAGC-CAGTCCCTGT 468
Qy      120 lIleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGl 140
Db      469 GATTGAGATAAAGTGTGCAATCGCTATGAGTTCTTGAATGGAAGAGTCCAATCCACCGA 528
Qy      140 uLeuCysAlaGlyHisLeuAlaGlyGlyThrAsp 151
Db      529 ACTCTGTGTGGGCATTTGGCCGGAGCACTGAC 562

```

## RESULT 6

N91337

LOCUS

DEFINITION

zaf4h05.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone

(IMAGE:292569 5', similar to gb:X05199 PLASMINOGEN PRECURSOR

(HUMAN); mRNA sequence.

N91337

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1995)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seg primer: mob.REGA+ET

High quality sequence stop: 296.

Location/Qualifiers

1..475

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:3800346"

/db\_xref="taxon:9606"

## FEATURES

source



COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongseung@mail.kribb.re.kr  
Plate: 41 row: G column: 11  
High quality sequence stop: 543.  
Location/Qualifiers  
1. 543

FEATURES  
source  
1. 543  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clones="L17N670205n1-41-G11"  
/sex="F"  
/lab host="Top10P."  
/clone lib="L17N670205n1"  
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI;  
Site 2: NotI; The library was contributed by the Soares  
laboratory and it was constructed as described by Bonaldo,  
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research  
6(9): 791-806. RNA was prepared from harvested cell  
culture."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.46e-92 Length: 543  
Score: 104.00 Matches: 104  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.24% Indels: 0  
DB: 6 Gaps: 0

US-09-992-095B-54 (1-207) x CB161910 (1-543)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20  
Db 230 ATGCACCTTCTGTGAGGACCTTGTATATCCCAAGAGTGGGTGTGACTGTGCCCCACTGC 289

Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40  
Db 290 TTGGAGAAGTCCCAAGGCTTCTATCTACAGGTCTATCTGCTGGTGCACCAAGAGTG 349

Qy 41 AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60  
Db 350 AATCTCGAACCGCATGTTTCAGGAATAAGAGTGTAGGCTGTCTTGGAGCCACACGA 409

Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspLysValIlePro 80  
Db 410 AAGATATTGCTTGTCTTAAAGCTAAGCAGTCTCTCCGCTCATCTGACAAAGTAAATCCCA 469

Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100  
Db 470 CTTGTCTGCCATCCCAAAATTATGTGCTGCTGACCGGACCGAATGTTTCATCATCTGGC 529

Qy 101 TrpGlyGluThr 104  
Db 530 TGGGAGAAACC 541

RESULT 9  
H60805  
LOCUS  
DEFINITION Yr45g09.r1 Soares fetal liver spleen INFLS Homo sapiens CDNA clone  
IMAGE:208288 5' similar to gb:X05199 PLASMINOGEN PRECURSOR  
(HUMAN); mRNA sequence.  
ACCESSION H60805  
VERSION H60805.1 GI:1013637  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 481)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parks, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: set@watson.wustl.edu  
High quality sequence stops: 339  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: M13RP1  
High quality sequence stop: 339.  
Location/Qualifiers  
1. 481  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3777419"  
/db\_xref="taxon:9606"  
/clone="IMAGE:208288"  
/sex="male"  
/dev stage="20 week-post conception fetus"  
/lab host="DH10B (ampicillin resistant)"  
/clone lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGGAGATTAATTAAGATCTTTTCTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN  
Alignment Scores:  
Pred. No.: 1.29e-91 Length: 481  
Score: 103.00 Matches: 103  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 49.76% Indels: 0  
DB: 7 Gaps: 0

US-09-992-095B-54 (1-207) x H60805 (1-481)

Qy 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20  
Db 80 ATGCACCTTCTGTGAGGACCTTGTATATCCCAAGAGTGGGTGTGACTGTGCCCCACTGC 139

Qy 21 LeuGluIysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40  
Db 140 TTGGAGAAGTCCCAAGGCTTCTATCTTACAGGTCTATCTGCTGGTGCACCAAGAGTG 199

Qy 41 AsnLeuGluProHisValGlnGluValSerArgLeuPheLeuGluProThrArg 60  
Db 200 AATCTCGAACCGCATGTTTCAGGAATAAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 259

Qy 61 LysAspIleAlaLeuLeuIysLeuSerSerProAlaValIleThrAspLysValIlePro 80  
Db 260 AAGATATTGCTTGTCTTAAAGCTAAGCAGTCTCTCCGCTCATCTGACAAAGTAAATCCCA 319

Qy 81 AlaCysLeuProSerProAsnTyrlsValAlaAlaAspArgThrGluCysPheIleThrGly 100  
Db 320 GCTTGTCTGCCATCCCAAAATTATGTGCTGCTGACCGGACCGAATGTTTCATCATCTGGC 379

Qy 101 TrpGlyGlu 103

```

Db      380 TGGGGGAA 388
|||||
RESULT 10
AV661991      594 bp  mRNA  linear  EST 16-JAN-2002
LOCUS      AV661991 GLC Homo sapiens cDNA clone GLCGZG06 3', mRNA sequence.
DEFINITION      AV661991
ACCESSION      AV661991
VERSION        AV661991.1 GI:9883005
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 594)
AUTHORS        Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
Hu,G., Gu,J., Chen,Z. and Han,Z.
TITLE          Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
MEDLINE        21625106
PUBMED        11752456
COMMENT        Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES       source
1..594
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GLCGZG06"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="GLC"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Alignment Scores:
Pred. No.: 1..46e-88 Length: 594
Score: 100.00 Matches: 100
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.31% Indels: 0
DB: 1 Gaps: 0
US-09-992-095B-54 (1-207) x AV661991 (1-594)
QY 1 MetHiePheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20
Db 70 ATGCACCTCTGTGGAGGACCTTGATATCCAGAGTGGGTGTGACTGCTGCCACTGC 129
QY 21 LeuGluYsSerProArgProSerSerTyrlsValIleLeuGluYsHisGlnGluVal 40
Db 130 TTGGAGAAAGTCCCAAGGCTTCATCTCAAGGTCTATCTGGGTGCACACCAAGAGTG 189
QY 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Db 190 AATCTCGAACCGCATGTCTCAGGAATAGAGTGTCTAGGCTGTCTCTGGAGCCACACGA 249
QY 61 LysAspIleAlaLeuLeuYsLeuSerSerProAlaValIleThrAspLysValIlePro 80
Db 250 AAAGATATTGCTTAAAGCTAAAGCAGTCTCCGCTCATCTACTGACAAAGTAATCCCA 309

```

---

```

QY 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
Db 310 GCTTGCTGTGCTCCCAATTTATGTGTCGCGACCGACCGAATGTTTCATCACTGGG 369
|||||
RESULT 11
H37877
LOCUS      H37877
DEFINITION      YP57f06.r1 Soares fetal liver spleen INF1S Homo sapiens cDNA clone
IMAGE:191555 5', similar to gb:X05199 PLASMINOGEN PRECURSOR
(HUMAN);, mRNA sequence.
ACCESSION      H37877
VERSION        H37877.1 GI:907376
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 381)
AUTHORS        Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Woldmann,P. and
Wilson,R.
TITLE          The WashU-Merck EST Project
JOURNAL        Unpublished (1995)
COMMENT        Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 843
High quality sequence stops: 278
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 843 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 278.
FEATURES       source
1..381
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3761344"
/db_xref="taxon:9606"
/clone="IMAGE:191555"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INF1S"
/note="Organ: Liver and Spleen; Vector: p7T73D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACTGGAAGAAATTAATAAGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.: 9..41e-87 Length: 381
Score: 98.00 Matches: 98
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 47.34% Indels: 0
DB: 7 Gaps: 0
US-09-992-095B-54 (1-207) x H37877 (1-381)
QY 51 ValSerArgLeuPheLeuGluProThrArgLysAspIleAlaLeuLeuYsLeuSer 70
|||||

```





# Tumor Gene Index

## JOURNAL COMMENT

Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov

Seq primer: -400P from Gibco  
 High quality sequence stop: 425.

## FEATURES

### source

Location/Qualifiers  
 1..593  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3183596"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Kid11"  
 /note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323176-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

### Alignment Scores:

Pred. No.:	1,248-80	Length:	593
Score:	92.00	Matches:	92
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	44.44%	Indels:	0
DB:	2	Gaps:	0

US-09-992-095B-54 (1-207) x BE348267 (1-593)

Qy	116	AlaGlnLeuProValIleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArg	135
Db	503	GCCAGCTCCCTGTGATTGAGAAATAAGTGTGCAATCGTATGAGTTCTGAATGGAAGA	444
Qy	136	ValGlnSerThrGluLeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGly	155
Db	443	GTCCAATCCACCGAACTCTGTCTGGCAATTTGGCCGAGGCACTGACAGTTGCCAGGGT	384
Qy	156	AspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThr	175
Db	383	GACAGTGGAGGTCCTCTGTTTCTTCGAGAAGGACAAATACATTTTACAGAGTCACT	324
Qy	176	SerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyValTyrValArgValSerArg	195
Db	323	TCTTGGGGTCTTGGCTGTGCACGCCCAATAAGCCTGGTGTCTATGTTCTGTTTCAAGG	264
Qy	196	PheValThrTrpIleGluGlyValMetArgAsnAsn	207
Db	263	TTTGTACTTGGATTGAGGGAGTGATGAGAAATAAT	228

Search completed: October 27, 2004, 11:47:59  
 Job time : 3067 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 09:13:27 ; Search time 153 Seconds  
(without alignments)  
778.448 Million cell updates/sec

Title: US-09-992-095B-54  
Perfect score: 207  
Sequence: 1 MHFCGGTLLSPWVLTAAHC.....GVVVRVSRFVTWIEGVMRNN 207

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 1825181 seqs, 575374646 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 783451

Minimum DB seq length: 0  
Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Database : Uniprot 02: \*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	18.8	73	Q9TV90	Q9TV90 equus cabal
2	32	15.5	54	Q6JDI3	Q6JDI3 canis famil
3	14	6.8	176	Q6SLI2	Q6SLI2 canis famil
4	14	6.8	176	AAR19224	AAR19224 canis fam
5	13	6.3	87	Q9CQ78	Q9CQ78 m mus muscu
6	13	6.3	176	Q91V08	Q91V08 mus musculu
7	12	5.8	126	Q7M325	Q7M325 aus scrofa
8	12	5.8	164	Q9DC82	Q9DC82 mus musculu
9	12	5.8	165	TRY3_LUCU	P35043 lucilia cup
10	11	5.3	119	Q9NR68	Q9NR68 homo sapien
11	11	5.3	149	Q25237	Q25237 lucilia cup
12	11	5.3	150	Q06784	Q06784 haematobia
13	11	5.3	175	Q6PLJ9	Q6PLJ9 aquilla ora
14	11	5.3	175	Q9GN96	Q9GN96 chrysomya b
15	11	5.3	175	Q9GSL6	Q9GSL6 chrysomya b
16	11	5.3	175	Q9GSM0	Q9GSM0 chrysomya b
17	11	5.3	175	Q9GSM1	Q9GSM1 chrysomya b
18	11	5.3	175	Q9GSM4	Q9GSM4 chrysomya b
19	11	5.3	175	AAT09986	AAT09986 squilla o
20	11	5.3	176	Q8KSD7	Q8KSD7 mus musculu
21	11	5.3	182	Q6PLJ6	Q6PLJ6 penaeus jap
22	11	5.3	182	Q6PLJ8	Q6PLJ8 procamburus
23	11	5.3	182	AAT09987	AAT09987 procambar
24	11	5.3	182	AAT09989	AAT09989 penaeus j
25	11	5.3	183	Q6PLJ7	Q6PLJ7 fennetropena
26	11	5.3	183	AAT09988	AAT09988 fennetropena
27	11	5.3	200	Q924U6	Q924U6 mus musculu
28	10	4.8	50	Q9GTF6	Q9GTF6 culex quinq
29	10	4.8	85	Q8MVL1	Q8MVL1 boltenia vi
30	10	4.8	116	Q7JMX5	Q7JMX5 helicoverpa
31	10	4.8	117	Q9PUP3	Q9PUP3 bothrops ja

32	10	4.8	125	2	Q804G0	Q80490 sphoeroides
33	10	4.8	141	2	Q6GKZ6	Q6GKZ6 drosophila
34	10	4.8	142	2	Q8HYM3	Q8HYM3 felis silve
35	10	4.8	147	2	Q8HZD0	Q8HZD0 saguinus oe
36	10	4.8	149	2	Q6DTY8	Q6DTY8 hypophthalm
37	10	4.8	154	2	O18448	O18448 helicoverpa
38	10	4.8	155	2	Q9YIK4	Q9YIK4 anopheles g
39	10	4.8	159	2	Q8HZD1	Q8HZD1 macaca sp.
40	10	4.8	159	2	Q8HZD2	Q8HZD2 pongo pygma
41	10	4.8	159	2	Q8HZD3	Q8HZD3 gorilla gor
42	10	4.8	159	2	Q8HZD4	Q8HZD4 pan troglod
43	10	4.8	160	2	Q6XGZ1	Q6XGZ1 homo sapien
44	10	4.8	160	2	AAP70247	AAP70247 homo sapi
45	10	4.8	178	2	O93594	O93594 dicentrarch

ALIGNMENTS

RESULT 1

Q9TV90 PRELIMINARY; PRT; 73 AA.

AC Q9TV90; (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DE Plasminogen (Fragment).

GN Name:PLG;

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OC NCBI\_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99160468; PubMed=10051323;

RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;

RT "Comparative mapping of 18 equine type I genes assigned by somatic cell hybrid analysis."

RL Mamm. Genome 10:271-276(1999).

DR EMBL; AF097581; RAD25984.1; -

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0004508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001254; Peptidase S1.

DR	InterPro; IPR009003; Pept\_Ser\_Cys.										
DR	Pfam; PF00089; Trypsin; 1.										
DR	PROSITE; PS0240; TRYPsin\_DOM; 1.										
DR	PROSITE; PS00135; TRYPsin\_SER; 1.										
DR	KW Hydrolase; Protease; Serine protease.										
FT	NON\_TER 1 1										
FT	NON\_TER 73 73										
SQ	SEQUENCE 73 AA; 7826 MW; 6CCBBFBA93D07704 CRC64;										
Query Match 18.8%; Score 39; DB 2; Length 73;											
Best Local Similarity 100.0%; Pred. No. 3.7e-32;											
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;											
Qy	151	DSQCQSGGGLVCFKDKYKILQGVTSWGLGCAFPNKPVG	189								
Db	35	DSQCQSGGGLVCFKDKYKILQGVTSWGLGCAFPNKPVG	73								
RESULT 2											
Q6JDI3 PRELIMINARY; PRT; 54 AA.											
AC Q6JDI3											
DT 05-JUL-2004 (Tremblrel. 27, Created)											
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)											
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)											
DE Plasminogen (Fragment).											
GN Name:PLG;											
OS Canis familiaris (Dog).											
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											

```
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed15233990;
RA Housley D.J.E., Ritzert E., Venta P.J.;
RT "Comparative radiation hybrid map of canine chromosome 1 incorporating
RT SNP and indel polymorphisms.";
RL Genomics 84:248-264(2004).
DR EMBL; AY514750; AAT44581.1; --
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 54
SQ SEQUENCE 54 AA; 5865 MW; CEF792BED3P4281D CRC64;

Query Match 15.5%; Score 32; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 5.5e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 LAGGTDSCQDGGGGLVCFEKDKYILQGVTSW 177
DB 23 LAGGTDSCQDGGGGLVCFEKDKYILQGVTSW 54

RESULT 3
Q6SL2 PRELIMINARY; PRT; 176 AA.
AC Q6SL2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Urokinase-type plasminogen activator (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Spee B., Penning L.C., Rothuizen J.;
RT "Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RT -1- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY455801; AAR19224.1; --
GO; GO:0016301; F.kinase activity; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1a.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Kinase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 19624 MW; 6B5B70BF55FAE708 CRC64;

Query Match 6.8%; Score 14; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 TDSQGDGGGGLVLC 163
DB 157 TDSQGDGGGGLVLC 170

RESULT 4
AAR19224 PRELIMINARY; PRT; 176 AA.
AC AAR19224;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DE Urokinase-type plasminogen activator (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Spee B., Penning L.C., Rothuizen J.;
RT "Differential gene expression of regenerative and fibrotic pathways in
RT canine hepatic portosystemic shunt and portal vein hypoplasia.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY455801; AAR19224.1; --
KW Kinase.
FT NON_TER 1
FT NON_TER 176
SQ SEQUENCE 176 AA; 19624 MW; 6B5B70BF55FAE708 CRC64;

Query Match 6.8%; Score 14; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 8.5e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 TDSQGDGGGGLVLC 163
DB 157 TDSQGDGGGGLVLC 170

RESULT 5
Q9CQ78 PRELIMINARY; PRT; 87 AA.
AC Q9CQ78;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Mus musculus adult male tongue cDNA, RIKEN full-length enriched
DE library, clone:2310081E03 product:protease, serine, 20, full insert
DE sequence (Mus musculus adult male tongue cDNA, RIKEN full-length
DE enriched library, clone:2310021N04 product:protease, serine, 20, full
DE insert sequence) (Fragment).
GN Name=2310015I08Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA The FANTOM Consortium.
RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
```





DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Trypsin alpha-3 (EC 3.4.21.4) (Fragment).  
 OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 CC Calliphoridae; Lucilia.  
 OX NCBI\_TaxID=7375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9502118; PubMed=7894748;  
 RA Casu R.E., Jarney J.M., Elvin C.M., Eisemann C.H.; Eisemann C.H.;  
 RT "Isolation of a trypsin-like serine protease gene family from the  
 RT sheep blowfly Lucilia cuprina."  
 RL Insect Mol. Biol. 3:159-170(1994).  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg|-Xaa, Lys|-Xaa.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L15632; AAA65931.1; --  
 DR HSP; P00763; IDPO.  
 DR MEROPS; S01.112; --  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS0134; TRYPSIN\_HIS; PARTIAL.  
 DR PROSITE; PS0135; TRYPSIN\_SER; 1.  
 DR Hydrolase; Multiogene family; Serine protease.  
 FT NON\_TER 1  
 FT ACT\_SITE 26 Charge relay system (By similarity).  
 FT ACT\_SITE 119 Charge relay system (By similarity).  
 FT DISULFID 89 106 By similarity.  
 FT DISULFID 115 139 By similarity.  
 FT SITE 113 113 Required for specificity (By similarity).  
 SQ SEQUENCE 165 AA; 18569 MW; 26160B1A9F90FCD CRC64;  
 Query Match 5.8%; Score 12; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.00097;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 151 DSCQSGSGGPLV 162  
 DB 113 DSCQSGSGGPLV 124  
 RESULT 10  
 ID Q9NR68 PRELIMINARY; PRT; 119 AA.  
 AC Q9NR68;  
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Serine protease kallikrein/ovasin/neuropsin type 3.  
 GN Name=KLK8;  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21206420; PubMed=11309326;  
 RA Magklara A., Scorilas A., Katsaros D., Massobrio M., Yousef G.M.,

RA Fracchioli S., Danese S., Diamandis E.P.;  
 RT "The human KLK8 (neuropain/ovasin) gene: identification of two novel  
 RT splice variants and its prognostic value in ovarian cancer."  
 RL Clin. Cancer Res. 7:806-811(2001).  
 DR EMBL; AF251125; AAF79144.1; --  
 DR HSP; P00760; IEZX.  
 DR GO; GO:0008233; F:trypsin activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS0135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 SQ SEQUENCE 119 AA; 12718 MW; 2FD8164DF1641PFF CRC64;  
 Query Match 5.3%; Score 11; DB 2; Length 119;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 153 CQDSDGGPLVC 163  
 DB 67 CQDSDGGPLVC 77  
 RESULT 11  
 ID Q25237 PRELIMINARY; PRT; 149 AA.  
 AC Q25237;  
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
 DE Serine proteinase (Fragment).  
 OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 CC Calliphoridae; Lucilia.  
 OX NCBI\_TaxID=7375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Long Pocket Laboratory culture; TISSUE=Total organism;  
 RX MEDLINE=95079073; PubMed=7987520;  
 RA Elvin C.M., Vuocolo T., Smith W.J., Eisemann C.H., Riddles P.W.;  
 RT "An estimate of the number of serine protease genes expressed in sheep  
 RT blowfly larvae (Lucilia cuprina)."  
 RL Insect Mol. Biol. 3:105-115(1994).  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 DR EMBL; U07693; AAA17385.1; --  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS0134; TRYPSIN\_HIS; UNKNOWN 1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 1  
 FT NON\_TER 149 149  
 SQ SEQUENCE 149 AA; 15212 MW; 95953C3945C317CD CRC64;  
 Query Match 5.3%; Score 11; DB 2; Length 149;  
 Best Local Similarity 100.0%; Pred. No. 0.0098;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 151 DSCQSDGGGPL 161  
 DB 139 DSCQSDGGGPL 149

```

RESULT 12
Q06784
ID Q06784 PRELIMINARY; PRT; 150 AA.
AC Q06784;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serine protease (Fragment).
OS Haematobia irritans (Horn fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;
OC Muscidae; Haematobia.
OX NCBI_TaxID=7368;
RN [1]
RP SEQUENCE FROM N.A. PubMed=8341258;
RX MEDLINE=93341451;
RA Elvin C.M., Whan V.A., Riddles P.W.;
RT "A family of serine protease genes expressed in adult buffalo fly
  (Haematobia irritans exiguus)."
RL Mol. Genet. 240:132-139(1993).
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; Z22567; CA80289.1; -.
DR PIR; S35208; S35208.
DR MEROPS; S01.110; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1 150
FT NON_TER 150 150
SQ SEQUENCE 150 AA; 16047 MW; 290D3054687FF37F CRC64;

Query Match 5.3%; Score 11; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPL 161
DB 140 DSCQDSDGGPL 150

RESULT 13
Q6PLJ9
ID Q6PLJ9 PRELIMINARY; PRT; 175 AA.
AC Q6PLJ9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Trypsin (Fragment).
OS Squilla oratoria.
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Hoplocarida; Stomatopoda; Squillidae; Squilla.
OX NCBI_TaxID=274635;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang N., Jiang G., Li N., Li L., Zhang J., Ma X., Chen Q.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY596941; AAT0986.1; -.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
```

```

DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1 175
FT NON_TER 175 175
SQ SEQUENCE 175 AA; 18536 MW; 671EEABD5D1FF20A CRC64;

Query Match 5.3%; Score 11; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPL 161
DB 152 DSCQDSDGGPL 162

RESULT 14
Q9GN96
ID Q9GN96 PRELIMINARY; PRT; 175 AA.
AC Q9GN96;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Serine protease K4.1/F1R2 (Serine protease K6.1/F1R1) (Serine protease
  K6.2/F1R2) (Fragment).
OS Chrysomya bezziana (Old world screwworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Chrysomya.
OX NCBI_TaxID=69364;
RN [1]
RP SEQUENCE FROM N.A. PubMed=11520682;
RX MEDLINE=21411890;
RA Muharsini S., Dalrymple B., Vuocolo T., Hamilton S., Willadsen P.,
  Wilfjels G.;
RT "Biochemical and molecular characterization of serine proteases from
  larvae of Chrysomya bezziana, the Old World Screwworm fly."
RL Insect Biochem. Mol. Biol. 31:1029-1040(2001).
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF302489; AAG30250.1; -.
DR EMBL; AF302487; AAG30248.1; -.
DR EMBL; AF302480; AAG30241.1; -.
DR HSP; P00761; 1BPT.
DR MEROPS; S01.110; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1 175
FT NON_TER 175 175
SQ SEQUENCE 175 AA; 18166 MW; 9DF0C459D83DC173 CRC64;

Query Match 5.3%; Score 11; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQDSDGGPL 161
DB 165 DSCQDSDGGPL 175
```

Thu Oct 28 07:18:45 2004

```
RESULT 15
Q9GSL6 PRELIMINARY; PRT; 175 AA.
ID Q9GSL6
AC Q9GSL6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serine protease K16/FIR2 (Fragment).
OS Chrysomya bezziana (Old world screwworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Calliphoridae; Chrysomya.
OX NCBI_TaxID=69364;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21411890; PubMed=11520682;
RA Muharsini S., Dalrymple B., Vuocolo T., Hamilton S., Willadsen P.,
RA Wjffels G.;
RT "Biochemical and molecular characterization of serine proteases from
RT larvae of Chrysomya bezziana, the Old World Screwworm fly.";
RL Insect Biochem. Mol. Biol. 31:1029-1040(2001).
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF302486; AAG30247.1; -.
DR MEROPS; S01.110; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 175
SQ SEQUENCE 175 AA; 18200 MW; 0AA9F30746F8FD3D CRC64;

Query Match 5.3%; Score 11; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 DSCQGDGGGPL 161
Db 165 DSCQGDGGGPL 175
|||||
|||||
```

Search completed: October 27, 2004, 09:31:57  
Job time : 155 secs

**This Page Blank (uspio)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 27, 2004, 07:37:41 ; Search time 89 Seconds  
(without alignments)  
5006.088 Million cell updates/sec

Title: US-09-992-095b-53\_COPY\_1044\_1664

Perfect score: 207

Sequence: 1 atgcactctgtggaggcac.....agggagtgtgagaaataat 621

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2559286

Minimum DB seq length: 0

Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DBV=xlh  
-Q=/cgn2\_1/USPTO.spool/US0992095/runat\_26102004\_085834\_10700/app.query.fasta\_1.775  
-DB=A\_Geneseq\_23Sep04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pcg -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=207  
-USER=US0992095 @CGN\_1\_1\_101 @runat\_26102004\_085834\_10700 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_23Sep04: \*  
1: Geneseq1980s: \*  
2: Geneseq1990s: \*  
3: Geneseq2000s: \*  
4: Geneseq2001s: \*  
5: Geneseq2002s: \*  
6: Geneseq2003as: \*  
7: Geneseq2003bs: \*  
8: Geneseq2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207	100.0	207	6	ABR48479 Human Pla
C 2	102	50.5	175	4	Aau32138 Novel hum
C 3	102	50.5	175	4	Aau32138 Novel hum
C 4	102	50.5	175	4	Aau32131 Novel hum
5	26	12.6	47	7	Adh34460 Plasmin,
6	18	8.7	95	4	Abg01924 Novel hum
7	17	8.2	84	5	Abp02684 Human ORF
8	14	6.8	118	4	Aab71662 Human col
9	13	6.3	13	8	Ado71577 Amino aci
10	13	6.3	14	5	Aae17240 Human tra

11	13	6.3	14	5	AAE18999	Human mat
12	13	6.3	18	4	ABBS2232	Human AFI
13	13	6.3	18	4	ABBS2105	Human AFI
14	13	6.3	18	6	ABR58948	Alzheimer
15	13	6.3	18	8	ADN31742	Human Alz
16	13	6.3	24	5	AAE17246	Serine pr
17	13	6.3	24	5	AAE19007	Human mat
18	13	6.3	35	5	AAE17239	Human tra
19	13	6.3	35	5	AAE19008	Human mat
20	13	6.3	35	5	AAE19008	Human tra
21	13	6.3	141	7	ADG75740	Human pro
22	13	6.3	149	4	AAU68926	Human pro
23	13	6.3	149	4	AAU68926	Protease
24	13	6.3	149	8	ADH78455	Human pro
25	13	6.3	156	6	ABR41514	Human DIT
26	13	6.3	157	7	AAU68929	Human pro
27	13	6.3	157	7	AAU68929	Protease
28	13	6.3	157	8	ADH78458	Human TAD
29	13	6.3	181	8	ADI16332	Human pro
30	13	6.3	191	6	ABG75787	Serine pr
31	13	6.3	193	4	AAU68929	Human pro
32	13	6.3	194	8	ADM79155	Human del
33	13	6.3	199	7	ADG42737	Human bet
34	13	6.3	199	7	ADJ55806	peptide h
35	13	6.3	199	8	ADM76644	Human NOV
36	12	5.8	12	3	AAU72111	Peptide f
37	12	5.8	12	4	AAU72111	Human col
38	12	5.8	12	5	ABG30800	Human ser
39	12	5.8	12	5	AAE17931	Human Gen
40	12	5.8	159	4	AAU68930	Human pro
41	12	5.8	159	7	ADH13375	Protease
42	12	5.8	159	8	ADH78459	Human TMP
43	11	5.3	11	8	ADO71578	Amino aci
44	11	5.3	16	3	AAE23914	Porcine v
45	11	5.3	30	3	AAE08953	Human sec

ALIGNMENTS

RESULT 1  
ABR48479  
ID ABR48479 standard; protein; 207 AA.  
XX  
AC ABR48479;  
XX  
DT 13-JUN-2003 (first entry)  
XX  
DE Human Plasminute.  
XX  
KW Human; GENSET; therapeutic; therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200294864-A2.  
XX  
PD 28-NOV-2002.  
XX  
PF 06-AUG-2001; 2001WO-1B001715.  
XX  
PR 25-MAY-2001; 2001US-0293574P.  
PR 15-JUN-2001; 2001US-0298698P.  
PR 29-JUN-2001; 2001US-0302277P.  
PR 13-JUL-2001; 2001US-0305456P.  
XX  
XX (GEST ) GENSET.  
PA ABR48479 Human Pla  
PI Aau32138 Novel hum  
PI Aau32138 Novel hum  
XX Adh34460 Plasmin,  
XX Abg01924 Novel hum  
XX Abp02684 Human ORF  
XX Aab71662 Human col  
XX Ado71577 Amino aci  
PT Aae17240 Human tra

New GENSET polynucleotides and polypeptides, useful for preparing a composition for treating GENSET-related disorders and as reagents in

PT assays to quantitatively determined levels of GENSET expression in  
PT biological samples.

PS Claim 2; Page 447-448; 505pp; English.

XX The present invention relates to novel human GENSET coding sequences  
CC (ACC51060-ACC51115) and proteins (ABR48453-ABR48508). The GENSET  
CC sequences are useful for preparing a composition for treating GENSET-  
CC related disorders. They can also be used as markers for tissues in which  
CC the corresponding protein is preferentially expressed, as molecular  
CC weight markers on Southern gels, as chromosome markers or tags to  
CC identify chromosomes, and as reagents in assays to quantitatively  
CC determined levels of GENSET expression in biological samples

XX Sequence 207 AA;

Alignment Scores:  
Pred. No.: 2,788-185 Length: 207  
Score: 207.00 Matches: 207  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x ABR48479 (1-207)

QY 1 ATGCACCTTCTGTGGAGGACCTTGTATATCCCGAGAGTGGGTGTGACTGTGCTGCCACTGC 60  
DB 1 MethHisPheCysGlyGlyThrLeuLeuSerProGluTyrValLeuThrAlaAlaHisCys 20  
QY 61 TTGGAGAGTCCCGCCAGGCTTCATCTACAAAGTTCCTCGGTGGTGCACACCAAGATG 120  
DB 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40  
QY 121 ATTCGACCCCATGTTCCAGAAATAGAGTGTCTAGCTGTCTTGGAGCCACACGA 180  
DB 41 AenLeuGluProHisValGlnGluLeuValSerArgLeuPheLeuGluProThrArg 60  
QY 181 AAGATATTCCTTGTGAAGTAAAGCAAGTCTCGCTCATCATCTGACAAAGTATCCCA 240  
DB 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80  
QY 241 GCTGTCTGCCATCCCGAAATATGTGTGCTGACCGGACCGAATGTTTCACTGCGC 300  
DB 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100  
QY 301 TGGGAGAAACCAAGGTACTTTTGGAGTGGCTTCTCAAGGAGCCAGCTCCCTGTG 360  
DB 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120  
QY 361 ATTGAGAATAAAGTGTGCAATCGCTATGAGTTTCTCAATGGAAGAGTCCATCCACCGAA 420  
DB 121 IleGluAenLysValCysAsnArgTyrGluPheLeuAenGlyArgValGlnSerThrGlu 140  
QY 421 CTCTGTCTGGGATTTGGCCGGAGGACTGACAGTTCGCGAGGTGACAGTGGAGTCT 480  
DB 141 LeuCysAlaGlyHisLeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160  
QY 481 CTGGTTTGTCTCGAGAGGACAAATACATTTTACAGGAGTCACTTCTTGGGGTCTTGGC 540  
DB 161 LeuValCysPheGluLysAspLysIleLeuGlnGlyValThrSerTrpGlyLeuGly 180  
QY 541 TGTGACGCGCCCAATAAGCTGTGTCTATGTTTCGTGTTTCAAGGTTTGTACTTGGATT 600  
DB 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200  
QY 601 GAGGAGTGTATGAGAAATAT 621  
DB 201 GluGlyValMetArgAsnAen 207

RESULT 2  
AAU32138  
ID AAU32138 standard; protein; 175 AA.

XX AAU32138;  
AC  
XX 18-DEC-2001 (first entry)  
DT  
XX Novel human secreted protein #2629.  
DE  
XX Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.  
XX Homo sapiens.  
OS  
XX WO200179449-A2.  
FN  
XX 25-OCT-2001.  
PD  
XX 16-APR-2001; 2001WO-US008656.  
PP  
XX 18-APR-2000; 2000US-00552929.  
PR  
XX 26-JAN-2001; 2001US-00770160.  
XX (HYSE-) HYSEQ INC.  
PA  
XX Tang YT, Liu C, Drmanac RT;  
PI  
XX WPI; 2001-611725/70.  
DR  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy.  
PT  
XX Claim 20; Page 561; 765pp; English.  
PS  
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising the  
CC nucleic acids encoding the polypeptides and cells genetically engineered  
CC to express them are also useful for producing the proteins. The proteins  
CC are useful in genetic vaccination, testing and therapy, and can be used  
CC as nutritional supplements. They may be used to increase stem cell  
CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
CC and/or nerve tissue growth or regeneration; immune suppression and/or  
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
CC AAU29510-AAU33304 represent the amino acid sequences of novel human  
CC secreted proteins of the invention

XX Sequence 175 AA;

Alignment Scores:  
Pred. No.: 1,18-86 Length: 175  
Score: 102.00 Matches: 102  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 50.50% Indels: 0  
DB: 4 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x AAU32138 (1-175)

QY 510 AATGATTGTGCTTCTCGAAGCAAAACAGAGGACCTCCACTGTCCCTGGCACTGTC 451  
DB 1 AsnValPheValLeuLeuGluAlaAsnGlnArgThrSerThrValThrLeuAlaThrVal 20  
QY 450 AGTGCCTCCGCCAAATGCCCCAGCACAGAGTTCGTTGGATTGGACTCTTCCATTCAGAA 391  
DB 21 SerAlaSerGlyGlnMetProSerThrGluPheGlyLeuAspSerSerIleGlnLys 40  
QY 390 CTCATAGCGATTGCACACTTTTATTCTCAATCACAGGAGCTGGGCTTCTTTGAGAGGCC 331



CC for use in treatment of a pathology related to aberrant expression or  
 CC physiological interactions of the polypeptide. Vectors comprising the  
 CC nucleic acids encoding the polypeptides and cells genetically engineered  
 CC to express them are also useful for producing the proteins. The proteins  
 CC are useful in genetic vaccination, testing and therapy, and can be used  
 CC as nutritional supplements. They may be used to increase stem cell  
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon  
 CC and/or nerve tissue growth or regeneration; immune suppression and/or  
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.  
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human  
 CC secreted proteins of the invention  
 XX  
 SQ Sequence 175 AA;

Alignment Scores:  
 Pred. No.: 1,1e-86 Length: 175  
 Score: 102.00 Matches: 102  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 50.50% Indels: 0  
 DB: 4 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x AAU32131 (1-175)

QY 510 AATGATTGCTCTTCGAGCAACACAGAGACCTCCACTGTCACCTGGCACTGTC 451  
 Db 1 AenValPheValLeuLeuGluAlaAsnGlnArgThrSerThrValThrLeuAlaThrVal 20  
 QY 450 ACTGCTCCGGCCAAATGCCAGCACAGATTCGGTGGATTGGACTTCCATTCCAGAAA 391  
 Db 21 SerAlaSerGlyGlnMetProSerThrGluPheGlyGlyLeuAspSerSerIleGlnLys 40  
 QY 390 CTCATAGCGATTGCACTATTTATCTCAATCACAGGAGCTGGGCTTCTTCAGAGGCC 331  
 Db 41 LeuIleAlaIleAlaHisPheIleLeuAsnHisArgGluLeuGlyPheLeuGluLysAla 60  
 QY 330 AGCTCCAAAGTACTTGGGTTTCTCCAGCCAGTGTATGAAACATTGGTCCGTCAGC 271  
 Db 61 SerSerLysSerThrLeuGlyPheSerProAlaSerAspGluThrPheGlyProValSer 80  
 QY 270 GACCACATATTTGGGGATGGCAGACAAGCTGGGATTACTTTGTTCAGTGTATGACGCGCAGG 211  
 Db 81 AspHisIleIleIleTrpGlyTrpGlnThrSerIlePheValSerAspAspGlyArg 100  
 QY 210 ACTGCT 205  
 Db 101 ThrAla 102

RESULT 5  
 ADH34460  
 ID ADH34460 standard; protein; 47 AA.  
 XX  
 AC ADH34460;

11-MAR-2004 (first entry)

Plasmin, a meginin ligand useful for treating renal failure.

Plasmin; meginin ligand; serpin; protease inhibitor;  
 Glomerular mesangial cell; drug screening;  
 Glomerular mesangial inflammation; renal failure.

Unidentified.

WO2003066089-A1.

14-AUG-2003.

07-FEB-2003; 2003WO-JP001316.

08-FEB-2002; 2002JP-00033164.

(KURO/) KUROKAWA K.

PA (MIYA/) MIYATA T.  
 XX Kurokawa K, Miyata T;  
 PI WPI; 2003-646264/61.  
 DR  
 XX Megsin ligand containing plasmin and/or trypsin for treatment of  
 PT glomerular mesangial inflammation.  
 XX  
 PS Disclosure; Page 12; 63pp; Japanese.

The invention relates to meginin ligands comprising plasmin and/or trypsin  
 as the active component. Megsin is a member of the serpin family of  
 protease inhibitors, binding to the serine proteases plasmin and trypsin,  
 thereby inhibiting their activity. Megsin is expressed in glomerular  
 mesangial cells and its overexpression is a cause of renal failure. The  
 invention also relates to a medical composition for treatment and  
 prevention of glomerular mesangial inflammation; a method of screening  
 for compounds that inhibit binding of meginin to meginin ligands; and the  
 compounds thus identified. The meginin ligands can be used in the  
 treatment and prevention of glomerular mesangial inflammation and renal  
 failure. The present sequence represents plasmin, a meginin ligand that  
 may be used in the invention.

SQ Sequence 47 AA;

Alignment Scores:  
 Pred. No.: 3,32e-15 Length: 47  
 Score: 26.00 Matches: 26  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.56% Indels: 0  
 DB: 7 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x ADH34460 (1-47)

QY 1 ATGCACATCTCTGAGGACCTTGATATCCCGAGTGGTGTGACTGTGCCACTGC 60  
 Db 1 MethHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20

QY 61 TTGGAGAGTCCCAAGG 78

Db 21 LeuGluLysSerProArg 26

RESULT 6

ABG01924

ID ABG01924 standard; protein; 95 AA.

XX AC ABG01924;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #1915.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

DR N-PSDB; AAS66111.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 32283; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG10377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 95 AA;

Alignment Scores:  
Pred. No.: 9.23e-08 Length: 95  
Score: 18.00 Matches: 18  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.70% Indels: 0  
DB: 4 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x ABG01924 (1-95)

QY 76 AGCGCTTCATCCTACAGGTCATCCCTGGTGACACCAAGTGAATCTCGAA 129  
Db 15 ArgProSerTyrIysValIleLeuGlyAlaHisGlnGluValAsnLeuGlu 32

RESULT 7  
ABP02684  
ID ABP02684 standard; protein; 84 AA.  
AC  
AC ABP02684;  
XX  
XX 24-JUN-2002 (first entry)  
XX  
XX Human ORFX protein sequence SEQ ID NO:5350.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;  
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;  
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;  
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;  
KW hypertension; hypothyroidism; cholesterol ester storage disease;  
KW immune deficiency; immune disorder; infectious disease;  
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;  
KW myasthenia gravis.

XX Homo sapiens.  
OS  
XX WO200192523-A2.  
PN  
XX  
XX 06-DEC-2001.  
PD  
XX  
XX 29-MAY-2001; 2001WO-US010836.  
PF

XX  
PR 30-MAY-2000; 2000US-0206132P.  
PR 29-AUG-2000; 2000US-0228716P.  
XX  
XX (CUPA-) CUPAGEN CORP.  
PA  
XX Shinkets RA, Leach MD;  
PI  
XX WPI; 2002-106308/14.  
XX N-PSDB; ABN18436.  
DR  
XX Novel human polypeptides and polynucleotides useful for diagnosing,  
PT preventing and treating cardiovascular disease, neurodegenerative,  
PT hyperproliferative disorders and autoimmune disorders.  
XX  
XX Disclosure; SEQ ID NO 5350; 1037pp; English.  
PS  
XX The present invention describes substantially purified human proteins  
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1  
CC in the specification). ABN15762 to ABN27252 encode the human ORFX  
CC proteins given in ABP0010 to ABP11500. ORFX proteins are useful for  
CC treating or preventing a pathology associated with an ORFX-associated  
CC disorder in humans, and in the manufacture of a medicament for treating a  
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide  
CC sequences can be used in gene therapy. ORFX sequences can be used in the  
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage. N.B. The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 84 AA;

Alignment Scores:  
Pred. No.: 8.23e-07 Length: 84  
Score: 17.00 Matches: 17  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.21% Indels: 0  
DB: 5 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x ABP02684 (1-84)

QY 232 GTAATCCAGCTGTCTGTCATCCCAATATATGTCGTCACCGACC 282  
Db 51 ValIleProAlaCysLeuProSerProAenTyrValValAlaAspArgThr 67

RESULT 8  
AAB71662  
ID AAB71662 standard; protein; 118 AA.  
XX  
XX AAB71662;  
AC  
XX 10-MAY-2001 (first entry)  
DT  
XX Human colon associated protein #10.  
DE  
XX Human; colon; cancer; disease.  
KW  
XX Homo sapiens.  
OS  
XX WO200112781-A1.  
PN

XX PD 22-FEB-2001.  
XX PF 11-AUG-2000; 2000WO-US022157.  
XX PR 13-AUG-1999; 99US-0148680P.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Birse CE, Rosen CA;  
XX XX WPI; 2001-147551/15.  
XX PT Nucleic acids encoding 13 human colon cancer associated polypeptides,  
PT useful for preventing, diagnosing and/or treating e.g. cancers  
PT (especially colon cancer), Parkinson's disease and diabetic retinopathy.  
XX PS Claim 11; Page 321; 326pp; English.  
XX CC The present invention relates to 13 human colon cancer-associated  
CC proteins. These proteins and the nucleic acid encoding them may be used  
CC in the prevention, diagnosis and treatment of diseases associated with  
CC inappropriate colon cancer-associated protein expression  
XX SQ Sequence 118 AA;  
Alignment Scores: Length: 118  
Pred. No.: 0.000501 Matches: 14  
Score: 14.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 6.76% Gaps: 0  
DB: 4  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x AAB71662 (1-118)  
QY 448 ACTGACAGTTCACGGTCACAGTCGAGTCCTCTGTTGC 489  
Db 56 ThrAspSerCysGlnGlyAspSerGlyProLeuValCys 69  
RESULT 9  
AD071577  
ID AD071577 standard; peptide; 13 AA.  
AC AD071577;  
XX 26-AUG-2004 (first entry)  
DT Amino acid sequence of microplasma loop domain peptide #2.  
DE microplasma; loop domain; alpha2-antiplasma; factor D; blood clot;  
XX thrombotic disorder; peripheral arterial occlusion;  
KW deep venous thrombosis; occlusion; blood vessel.  
XX Unidentified.  
OS WO2004045558-A2.  
XX 03-JUN-2004.  
XX 18-NOV-2003; 2003WO-US037221.  
XX 18-NOV-2002; 2002US-0427152P.  
XX (HARD ) HARVARD COLLEGE.  
XX Reed GL;  
XX WPI; 2004-420534/39.  
XX Novel alpha2-antiplasma resistant microplasma polypeptide comprising  
PT heterologous loop domain sequence, useful for treating thrombotic  
PT disorders such as deep venous thrombosis.

XX PS Disclosure; Page 2; 30pp; English.  
XX CC The specification describes a microplasma polypeptide, comprising a  
CC heterologous loop domain sequence, where the polypeptide is resistant to  
CC alpha2-antiplasma inhibition compared to a wild-type microplasma. The  
CC heterologous loop domain comprises at least 4 or 10 consecutive amino  
CC acids of a factor D loop domain, and is present in microplasma loop 3,  
CC loop 5, loop 6 or loop 7. The microplasma polypeptide of the invention  
CC is useful for dissolving a blood clot. It is useful for treating a  
CC such as peripheral arterial occlusion, deep venous thrombosis or other  
CC disorder associated with occlusion of a blood vessel. AD071576-ADO71579  
CC represent peptides from microplasma loop 3, loop 5, loop 6 and loop 7,  
CC respectively. These peptides are replaced with factor D loop domain  
CC peptides to create microplasma polypeptides of the invention.  
XX SQ Sequence 13 AA;  
Alignment Scores: Length: 13  
Pred. No.: 0.00707 Matches: 13  
Score: 13.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 6.28% Gaps: 0  
DB: 8  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x AD071577 (1-13)  
QY 52 GCCCAGCTGCTGGAGAGTCCCAAGGCTTCATCTCTAC 90  
Db 1 AlaHisCysLeuGluLysSerProArgProSerSerTyr 13  
RESULT 10  
AAE17240  
ID AAE17240 standard; peptide; 14 AA.  
AC AAE17240;  
XX 18-APR-2002 (first entry)  
DT Human transmembrane serine protease-related type I fibronectin domain #1.  
DE Transmembrane serine protease; gene therapy; metastasis; tumour;  
XX chronic obstructive pulmonary disease; COPD; angiogenesis; inflammation;  
KW atherosclerosis; neurodegenerative disease; neuroprotective; cytostatic;  
KW pathogenic infection; antiinflammatory; antiarteriosclerotic;  
KW antibacterial; type I fibronectin domain.  
XX Unidentified.  
OS WO200196538-A2.  
XX 20-DEC-2001.  
XX 12-JUN-2001; 2001WO-EP0066618.  
XX 13-JUN-2000; 2000US-0211224P.  
XX 13-APR-2001; 2001US-0283353P.  
XX 16-APR-2001; 2001US-0283648P.  
XX (FARB ) BAYER AG.  
XX Xiao Y, Gedrich R;  
XX WPI; 2002-098065/13.  
XX Novel isolated polynucleotide encoding transmembrane serine protease  
PT polypeptide, for treating chronic obstructive pulmonary disease, tumor  
PT angiogenesis, inflammation, atherosclerosis and neurodegenerative  
PT disease.  
XX Disclosure; Fig 3; 120pp; English.

XX The present invention relates to an isolated polynucleotide encoding a  
CC transmembrane serine protease polypeptide. Transmembrane serine protease  
CC gene is useful in gene therapy. The invention also relates to a  
CC pharmaceutical composition which is useful for modulating the activity of  
CC transmembrane serine protease in a disease, such as chronic obstructive  
CC pulmonary disease (COPD), metastasis of malignant cells, tumour  
CC angiogenesis, inflammation, atherosclerosis, a neurodegenerative disease  
CC or pathogenic infection. Transmembrane serine protease is useful as a  
CC bait protein in a two-hybrid or three-hybrid assay. The polypeptide is  
CC useful for generating antibodies against it and in various assay systems.  
CC The present sequence is a human transmembrane serine protease-related  
CC type I fibronectin domain  
XX  
SQ Sequence 14 AA;

Alignment Scores: Length: 14  
Pred. No.: 0.00696 Matches: 13  
Score: 13.00 Conservativeness: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 6.28% Gaps: 0  
DB: 5

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x AA17240 (1-14)

QY 451 GACAGTTCAGGTCACAGTCAGGTCCTCTGTTGC 489  
|||||  
DB 2 AspSerCysGlnGlyAspSerGlyProLeuValCys 14

RESULT 11  
AAE18999  
ID AAE18999 standard; peptide; 14 AA.  
XX  
AC AAE18999;  
XX  
XX 21-MAY-2002 (first entry)  
XX  
DE Human matrilysin-like serine protease related fibronectin domain #1.  
XX  
KW Human; matrilysin-like serine protease; cardiovascular; antiinflammatory;  
KW cytostatic; chronic obstructive pulmonary disease; pulmonary; vulnary;  
KW virucide; nervous system disorder; extracellular matrix degradation;  
KW antimicrobial; neuroprotective; cancer; cardiovascular disorder; wound;  
KW adenocarcinoma; neurodegenerative disease; Alzheimer's disease; melanoma;  
KW Genstmann-Straussler Syndrome; Creutzfeldt-Jakob disease; scrapie;  
KW Parkinson's disease; enzyme; type I fibronectin domain.  
XX  
OS Unidentified.  
XX  
XX WO200208392-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 16-JUL-2001; 2001WO-EP008182.  
XX  
XX 25-JUL-2000; 2000US-0220807P.  
XX  
XX 02-APR-2001; 2001US-0280109P.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Xiao Y;  
XX  
XX WPI; 2002-195871/25.  
XX  
XX New matrilysin-like serine protease proteins and polynucleotides, useful  
XX for treating matrilysin-like serine protease dysfunction related  
XX diseases, e.g. cancer, pulmonary disease, wounds, inflammation or viral  
XX infections.  
XX  
XX Disclosure; Fig 9; 140pp; English.  
XX  
XX The patent discloses novel matrilysin-like serine protease proteins and

CC their corresponding polynucleotides. The invention further relates to  
CC reagents and methods of regulating human matrilysin-like serine protease  
CC activity. Pharmaceutical compositions comprising the reagents of the  
CC invention are useful for modulating the activity of a matrilysin-like  
CC serine protease in a disease. The reagents are also useful for treating  
CC or ameliorating matrilysin-like serine protease dysfunction related  
CC diseases such as cancer, chronic obstructive pulmonary disease, central  
CC or peripheral nervous system disorder and cardiovascular disorder. The  
CC human matrilysin-like serine protease gene provides a therapeutic target  
CC to decrease extracellular matrix degradation, in particular for treating  
CC or preventing metastatic cancer, e.g. adenocarcinoma, melanoma, cancers  
CC of the adrenal gland, bladder, bone breast, cervix, gall bladder, liver,  
CC lung, ovary, pancreas, prostate, testis or uterus. Sequences of the  
CC invention are also useful for treating neurodegenerative diseases (e.g.,  
CC Alzheimer's disease, Parkinson's disease), wounds, inflammation or viral  
CC infections. Matrilysin-like serine protease activity can be used to  
CC degrade, prion protein amyloid plaques of Genstmann-Straussler Syndrome,  
CC Creutzfeldt-Jakob disease and Scrapie. The present sequence is human  
CC matrilysin-like serine protease related type I fibronectin domain  
XX  
SQ Sequence 14 AA;

Alignment Scores: Length: 14  
Pred. No.: 0.00696 Matches: 13  
Score: 13.00 Conservativeness: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 6.28% Gaps: 0  
DB: 5

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x AA18999 (1-14)

QY 451 GACAGTTCAGGTCACAGTCAGGTCCTCTGTTGC 489  
|||||  
DB 2 AspSerCysGlnGlyAspSerGlyProLeuValCys 14

RESULT 12  
ABBS2232  
ID ABBS2232 standard; peptide; 18 AA.  
XX  
AC ABBS2232;  
XX  
XX 08-FEB-2002 (first entry)  
XX  
DE Human API-180 tryptic digest peptide #4.  
XX  
KW Human; neuroprotective; nootropic; gene therapy; vaccine;  
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;  
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;  
KW Expression Reference Protein Isoform; ERPI; proteolysis.  
XX  
OS Homo sapiens.  
XX  
XX WO200175454-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 03-APR-2001; 2001WO-US010908.  
XX  
XX 03-APR-2000; 2000US-0194504P.  
XX  
XX 28-NOV-2000; 2000US-0253647P.  
XX  
XX (OXFO-) OXFORD GLYSCSCIENCES UK LTD.  
XX  
XX (PFIZ ) PFIZER INC.  
XX  
XX Durham KL, Friedman DL, Herath HM, Kimmel LH, Parekh RB;  
XX Potter DM, Rohlf C, Silber BW, Stiger TR, Sunderland PT;  
XX Townsend RR, White F, Williams SA;  
XX WPI; 2001-639384/73.  
XX  
XX Screening for Alzheimer's disease in a mammal, by making two-dimensional  
XX array of a feature whose relative abundance correlates with disease, and  
XX PT

PT comparing with abundance of the feature in samples of healthy persons.  
XX Example; Page 31; 162pp; English.  
XX  
CC The invention relates to methods for the screening, diagnosis and  
CC prognosis of Alzheimer's disease. The methods involve the detection of  
CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-  
CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or  
CC plasma. The abundance of the AFs and APIs is then normalised to an  
CC Expression Reference Protein Isoform (ERPI) in order to determine whether  
CC a patient is suffering from, or has a predisposition to, Alzheimer's  
CC Disease. The relative abundance of the AFs and APIs correlates with the  
CC severity of Alzheimer's disease. The present sequence is a peptide  
CC produced from an API by proteolysis  
XX  
SQ Sequence 18 AA;  
  
Alignment Scores:  
Pred. No.: 0.00658 Length: 18  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x ABB52232 (1-18)  
QY 451 GACAGTGGCCAGGTGACAGTGGAGTCTCTGTTTGC 489  
|||||  
Db 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13  
  
RESULT 13  
ABB52105  
ID ABB52105 standard; peptide; 18 AA.  
XX  
AC ABB52105;  
XX  
DT 08-FEB-2002 (first entry)  
XX  
DE Human API-26 tryptic digest peptide #3.  
XX  
KW Human; neuroprotective; nootropic; gene therapy; vaccine;  
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;  
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;  
KW Expression Reference Protein Isoform; ERPI; proteolysis.  
XX  
OS Homo sapiens.  
XX  
PN WO200175454-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 03-APR-2001; 2001WO-US010908.  
XX  
PR 03-APR-2000; 2000US-0194504P.  
PR 28-NOV-2000; 2000US-0253647P.  
XX  
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
PA (PFIZ ) PFIZER INC.  
XX  
XX Durham KL, Friedman DL, Herath HMA, Kimmel LH, Parekh RB;  
PI Potter DM, Rohlf C, Silber BM, Stiger TR, Sunderland PT;  
PI Townsend RR, White F, Williams SA;  
XX  
WPI; 2001-639384/73.  
XX  
XX Screening for Alzheimer's disease in a mammal, by making two-dimensional  
XX array of a feature whose relative abundance correlates with disease, and  
PT comparing with abundance of the feature in samples of healthy persons.  
XX  
XX Example; Page 28; 162pp; English.  
XX  
CC The invention relates to methods for the screening, diagnosis and

CC prognosis of Alzheimer's disease. The methods involve the detection of  
CC Alzheimer's Disease-Associated Features (AFs) and Alzheimer's Disease-  
CC Associated Protein Isoforms (APIs) in cerebrospinal fluid, serum or  
CC plasma. The abundance of the AFs and APIs is then normalised to an  
CC Expression Reference Protein Isoform (ERPI) in order to determine whether  
CC a patient is suffering from, or has a predisposition to, Alzheimer's  
CC Disease. The relative abundance of the AFs and APIs correlates with the  
CC severity of Alzheimer's disease. The present sequence is a peptide  
CC produced from an API by proteolysis  
XX  
SQ Sequence 18 AA;  
  
Alignment Scores:  
Pred. No.: 0.00658 Length: 18  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x ABB52105 (1-18)  
QY 451 GACAGTGGCCAGGTGACAGTGGAGTCTCTGTTTGC 489  
|||||  
Db 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13  
  
RESULT 14  
ABR58948  
ID ABR58948 standard; peptide; 18 AA.  
XX  
AC ABR58948;  
XX  
DT 11-JUL-2003 (first entry)  
XX  
DE Alzheimer's Disease-associated protein isoform, API-26, SEQ ID 68.  
XX  
KW Nootropic; Neuroprotective; Alzheimer's disease; API; human;  
KW Alzheimer's Disease-associated protein isoform.  
XX  
OS Homo sapiens.  
XX  
PN WO2003028543-A2.  
XX  
PD 10-APR-2003.  
XX  
PF 03-OCT-2002; 2002WO-US031642.  
XX  
PR 03-OCT-2001; 2001US-0326708P.  
XX  
PA (PFIZ ) PFIZER PROD INC.  
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
XX  
XX Durham LM, Friedman DL, Herath HMA, Kimmel LH, Parekh RB;  
PI Potter DM, Rohlf C, Silber BM, Snyder PJ, Soares HD, Stiger TR;  
PI Sunderland PT, Townsend RR, White WF, Williams SA;  
XX  
WPI; 2003-371957/35.  
XX  
XX Screening or diagnosing of Alzheimer's disease (AD) determine the stage  
XX or severity of AD in a subject, comprises analyzing a test sample of body  
XX fluid from the subject by 2-dimensional electrophoresis.  
XX  
XX Claim 2; Page 41; 179pp; English.  
XX  
XX The present invention relates to methods for screening or diagnosing  
XX Alzheimer's disease (AD) to determine the stage or severity of AD in a  
XX subject, to identify subject at risk of developing AD, or to monitor the  
XX effect of therapy administered. The methods comprise analysing a test  
XX sample of body fluid by 2-dimensional electrophoresis to generate a 2-  
XX dimensional array of AD-associated features (AFs). The method  
XX alternatively comprises quantitatively detecting in a sample of body  
XX fluid from the subject, one or more AD-associated protein isoforms (APIs;  
XX ABR58710-ABR59184)



```
XX SQ Sequence 18 AA;
XX Alignment Scores:
XX Pred. No.: 0.00658 Length: 18
XX Score: 13.00 Matches: 13
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 6.28% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-992-095B-53_COPY_1044_1664 (1-621) x ABR58948 (1-18)
XX QY 451 GACAGTTCGACAGTGCAGGTCTCTGTTTGC 489
XX DB 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13
XX
XX RESULT 15
XX ID ADN31742
XX ID ADN31742 standard; peptide; 18 AA.
XX AC ADN31742;
XX DT 01-JUL-2004 (first entry)
XX DE Human Alzheimer's disease-API tryptic digest peptide - SEQ ID 68.
XX KW Alzheimer's disease; neurotropic; neuroprotective; cerebrospinal fluid;
XX KW CSF; Alzheimer's disease-associated protein isoform; API; tryptic digest;
XX KW human.
XX OS Homo sapiens.
XX PN EP1408333-A2.
XX PD 14-APR-2004.
XX PF 03-OCT-2002; 2002EP-00256893.
XX PR 03-OCT-2002; 2002US-0326708P.
XX PA (PFIZ ) PFIZER PROD INC.
XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX PI Durham LK, Friedman DL, Herath HMAK, Kimmel LH, Parekh RB;
XX PI Potter DM, Rohlf C, Silber BM, Snyder PJ, Soares HD, Stiger TR;
XX PI Sunderland PT, Townsend RR, White WF, Williams SA;
XX DR WPI; 2004-318939/30.
XX
XX Screening or diagnosis of Alzheimer's disease (AD) in subject,
XX determining stage or severity of AD, identifying subject at risk of
XX developing AD, or monitoring effect of therapy, by detecting Alzheimer's
XX disease-Associated Features.
XX Example; SEQ ID NO 68; 208pp; English.
XX
XX The invention relates to a novel method for screening or diagnosis of
XX Alzheimer's disease (AD) in a subject, determining the stage or severity
XX of AD, identifying a subject at risk of developing AD or monitoring the
XX effect of therapy administered to a subject having AD, by analysing body
XX fluid to generate a two-dimensional array of Alzheimer's disease-
XX associated features (AFs) such as AF-200, AF-201, AF-202, AF-203, AF-204,
XX AF-205, etc., and comparing the abundance of AFs with a control. The
XX method of the invention has neurotropic and neuroprotective applications
XX and may be useful for screening or diagnosis of Alzheimer's disease (AD)
XX in a subject, determining the stage or severity of AD in a subject,
XX identifying a subject at risk of developing AD or monitoring the effect
XX of therapy administered to a subject having AD. The body fluid is
XX cerebrospinal fluid (CSF). The current sequence is that of a human
XX Alzheimer's disease-associated protein isoform (API) tryptic digest
XX peptide of the invention.
```

```
SQ Sequence 18 AA;
Alignment Scores:
Pred. No.: 0.00658 Length: 18
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.28% Indels: 0
DB: Gaps: 0
US-09-992-095B-53_COPY_1044_1664 (1-621) x ADN31742 (1-18)
QY 451 GACAGTTCGACAGTGCAGGTCTCTGTTTGC 489
DB 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13
Search completed: October 27, 2004, 09:07:52
Job time : 96 secs
```

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 27, 2004, 08:37:16 ; Search time 20.5 Seconds  
(without alignments)  
4017.900 Million cell updates/sec

Title: US-09-992-095B-53\_COPY\_1044\_1664

Perfect score: 207

Sequence: 1 atgcactctgtggaggcac.....agggagtgatgagaataaat 621

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 478139 seqs, 66318000 residues

Word size: 1

Total number of hits satisfying chosen parameters: 632370

Minimum DB seq length: 0

Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09992095/runat\_26102004\_085836\_10739/app\_query.fasta\_1.775  
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEAPSIEZE=500 -MINLEN=0 -MAXLEN=207  
-USER=US09992095 @CGN\_1\_1 22\_orunat\_26102004\_085836\_10739 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile61.pcp.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	207	100.0	207	4	US-10-000-489-54
2	42	20.3	197	1	US-08-456-840-48
3	42	20.3	197	1	US-08-266-407A-48
4	42	20.3	197	2	US-08-892-544-48
5	13	6.3	23	4	US-09-879-792-22
6	13	6.3	149	3	US-09-518-046-20
7	13	6.3	157	3	US-09-518-046-23
8	12	5.8	151	4	US-09-270-767-33178
9	12	5.8	151	4	US-09-270-767-48395
10	12	5.8	159	3	US-09-518-046-24
11	11	5.3	12	4	US-09-879-792-16
12	11	5.3	84	3	US-08-906-769-99

13	11	5.3	84	3	US-08-906-616-99	Sequence 99, Appl
14	11	5.3	84	3	US-08-817-795-99	Sequence 99, Appl
15	11	5.3	84	3	US-08-639-075A-99	Sequence 99, Appl
16	11	5.3	84	3	US-09-012-431-99	Sequence 99, Appl
17	11	5.3	84	3	US-09-012-692-99	Sequence 99, Appl
18	11	5.3	84	3	US-08-906-613-99	Sequence 99, Appl
19	11	5.3	84	5	PCT-US95-14442A-99	Sequence 99, Appl
20	11	5.3	138	6	5200340-4	Patent No. 5200340
21	11	5.3	162	4	US-09-244-111-6	Sequence 6, Appl
22	11	5.3	207	4	US-09-244-111-4	Sequence 4, Appl
23	10	4.8	42	1	US-08-293-778-14	Sequence 14, Appl
24	10	4.8	151	3	US-09-518-046-21	Sequence 21, Appl
25	10	4.8	154	3	US-09-261-416-5	Sequence 5, Appl
26	10	4.8	156	3	US-09-261-416-6	Sequence 6, Appl
27	10	4.8	182	4	US-09-328-925-12	Sequence 12, Appl
28	10	4.8	200	3	US-09-008-271A-5	Sequence 5, Appl
29	9	4.3	9	1	US-07-819-361-5	Sequence 5, Appl
30	9	4.3	9	1	US-08-179-574-5	Sequence 5, Appl
31	9	4.3	9	3	US-09-518-046-58	Sequence 58, Appl
32	9	4.3	9	4	US-09-618-259-27	Sequence 27, Appl
33	9	4.3	9	5	PCT-US93-00325-5	Sequence 5, Appl
34	9	4.3	20	1	US-08-472-228A-19	Sequence 19, Appl
35	9	4.3	20	3	US-09-146-831-19	Sequence 19, Appl
36	9	4.3	20	5	PCT-US96-09303-19	Sequence 19, Appl
37	9	4.3	22	1	US-08-392-828C-31	Sequence 31, Appl
38	9	4.3	22	3	US-09-330-945-31	Sequence 31, Appl
39	9	4.3	23	4	US-09-879-792-15	Sequence 15, Appl
40	9	4.3	36	3	US-08-944-483-26	Sequence 26, Appl
41	9	4.3	55	4	US-08-843-076D-22	Sequence 22, Appl
42	9	4.3	86	1	US-08-485-455D-53	Sequence 53, Appl
43	9	4.3	86	2	US-08-482-130C-53	Sequence 53, Appl
44	9	4.3	86	2	US-08-484-211C-53	Sequence 53, Appl
45	9	4.3	86	3	US-08-906-769-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1  
US-10-000-489-54  
; Sequence 54, Application US/10000489  
; Patent No. 6794363  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91 US6 DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-54

Alignment Scores:  
Pred. No.: 3,846-189  
Score: 207.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Length: 207  
Matches: 207  
Conservative: 0  
Mismatch: 0

```

Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-10-000-489-54 (1-207)

Qy 1 ATGCACCTCTGTGGAGGACCTTGATATCCACAGAGTGGTGTGACCTGCCACTGC 60
Db 1 MetHisPheCysGlyGlyThrLeuIleSerProGluTrpValLeuThrAlaHisCys 20
Qy 61 TTGGAGAAAGTCCCAAGGCGCTTCATCCCTACAGGTTCATCCCTGGGTGCACCAAGAAGTG 120
Db 21 LeuGluLysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40
Qy 121 AATCTGGAACCCATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGA 180
Db 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
Qy 181 AAAGATATTGCTTCAAGCTAAGCTAGCAGTCTCTCGCTCATCACTGACAAAGTAAATCCA 240
Db 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
Qy 241 GCTTCTCTGCCATCCCAATTTATGTGCTGCTGACCGGACCGAATGTTTCATCACTGCG 300
Db 81 AlaCysLeuProSerProAsnTyrlsValValAlaAspArgThrGluCysPheIleThrGly 100
Qy 301 TGGGAGAAACCAAGGTACTTTTGGAGTGGCTTCTCAAGGAAGCCAGCTCCCTGTG 360
Db 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
Qy 361 ATTGAGATTAAGTGTCAATCGCTATGATGTTCTGTAATGGAAGAGTCCAAATCCACGAA 420
Db 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
Qy 421 CTCTCTGTGGCAATTTGGCGGAGGACCTGACAGTTGCCAGGTGCAGTGGAGGTCT 480
Db 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
Qy 481 CTGGTTTGTTCGAGAGGACAAATACATTTTACAGAGTCACTTCTTGGGTCTTGGC 540
Db 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
Qy 541 TGTGCACGCCCAATAAGCCTGGTCTATGTTCTGTTCAAGTTTGTACTTGGATT 600
Db 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
Qy 601 GAGGAGTGCATGAGAAATAAT 621
Db 201 GluGlyValMetArgAsnAsn 207

```

```

RESULT 2
US-08-456-840-48
; Sequence 48, Application US/08456840
; Patent No. 5597908
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5597908el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,840
; FILING DATE: 01-JUN-1995

```

```

; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 197 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; US-08-456-840-48

Alignment Scores:
Pred. No.: 1,02e-31 Length: 197
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.29% Indels: 0
DB: 1 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-08-456-840-48 (1-197)

Qy 445 GGCACCTGACAGTTCACAGGTGACAGTGGAGTCTCTGTTGCTTCGAGAGGACAAA 504
Db 141 GlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLysAspLys 160
Qy 505 TACATTTTACAGAGTCACTTCTTGGGTCTTGGCTGTGTCAGCCCAATTAAGCTGGT 564
Db 161 TyrIleLeuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGly 180
Qy 565 GTCTAT 570
Db 181 ValTyr 182

RESULT 3
US-08-266-407A-48
; Sequence 48, Application US/08266407A
; Patent No. 5786156
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5786156el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407A
; FILING DATE: 27-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:

```

NAME: Gormley, Mary E.  
REGISTRATION NUMBER: 34,409  
TELEPHONE: 301-258-5200  
TELEFAX: 301-977-0847  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-266-407A-48

Alignment Scores:  
Pred. No.: 1,02e-31 Length: 197  
Score: 42.00 Matches: 42  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.29% Indels: 0  
DB: 1 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-08-266-407A-48 (1-197)

QY 445 GGCACGTGACAGTGGCCAGGTCACAGTGGAGGTCTCTGGTTTGTTCGAGAAGGACAAA 504  
Db 141 GlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGlnLysAspLys 160

QY 505 TACATTTTACAGGAGTCACTCTTGGGGTCTTGGCTGTGCAGCCCAATAAGCCTGCT 564  
Db 161 TyrlleuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAnLysProGly 180

QY 565 GTCTAT 570  
Db 181 ValTyr 182

RESULT 4  
US-08-892-544-48  
Sequence 48, Application US/08892544  
Patent No. 5874544  
GENERAL INFORMATION:  
APPLICANT: Taddei-Peters, W. C.  
APPLICANT: Butler, Sandra M.  
TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
NUMBER OF SEQUENCES: 48  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akzo No. 5874544e1  
STREET: 1330 Piccard Drive  
CITY: Rockville  
STATE: Maryland  
COUNTRY: US  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/892,544  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/266,407  
FILING DATE: 27-JUN-1994  
APPLICATION NUMBER: US 08/172,461  
FILING DATE: 21-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Gormley, Mary E.  
REGISTRATION NUMBER: 34,409  
TELEPHONE: 301-258-5200  
TELEFAX: 301-977-0847

INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 197 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-892-544-48

Alignment Scores:  
Pred. No.: 1,02e-31 Length: 197  
Score: 42.00 Matches: 42  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 20.29% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-08-892-544-48 (1-197)

QY 445 GGCACGTGACAGTGGCCAGGTCACAGTGGAGGTCTCTGGTTTGTTCGAGAAGGACAAA 504  
Db 141 GlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGlnLysAspLys 160

QY 505 TACATTTTACAGGAGTCACTCTTGGGGTCTTGGCTGTGCAGCCCAATAAGCCTGCT 564  
Db 161 TyrlleuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAnLysProGly 180

QY 565 GTCTAT 570  
Db 181 ValTyr 182

RESULT 5  
US-09-879-792-22  
Sequence 22, Application US/09879792  
Patent No. 6734006  
GENERAL INFORMATION:  
APPLICANT: Xiao, Yonghong  
APPLICANT: Gedrich, Richard  
TITLE OF INVENTION: Regulation of Human Transmembrane Serine  
TITLE OF INVENTION: Protease  
FILE REFERENCE: 02973.00035  
CURRENT APPLICATION NUMBER: US/09/879,792  
CURRENT FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/211,224  
PRIOR FILING DATE: 2000-06-13  
PRIOR APPLICATION NUMBER: US 60/283,353  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: US 60/283,648  
PRIOR FILING DATE: 2001-04-16  
PRIOR APPLICATION NUMBER: PCT  
PRIOR FILING DATE: 2001-06-12  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 23  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: BLOCKS BL00134B  
US-09-879-792-22

Alignment Scores:  
Pred. No.: 0.000637 Length: 23  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 4 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-09-879-792-22 (1-23)

QY 451 GACAGTTGCCAGGTCACAGTGGAGGTCTCTGGTTTGC 489

Db 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13  
|||||  
RESULT 6  
US-09-518-046-20  
; Sequence 20, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 20  
; LENGTH: 149  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: protease domain of protease M (Prom)  
US-09-518-046-20  
Alignment Scores:  
Pred. No.: 0.000495 Length: 149  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 3 Gaps: 0  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-09-518-046-20 (1-149)  
Qy 451 GACAGTTGCCAGGGTGACAGTGGAGGTCCTCTGGTTGC 489  
|||||  
Db 137 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 149  
RESULT 7  
US-09-518-046-23  
; Sequence 23, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 23  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: protease domain of TADG-12 (TADG12)  
US-09-518-046-23  
Alignment Scores:  
Pred. No.: 0.000492 Length: 157  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 3 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-09-518-046-23 (1-157)  
Qy 451 GACAGTTGCCAGGGTGACAGTGGAGGTCCTCTGGTTGC 489  
|||||  
Db 145 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 157  
RESULT 8  
US-09-270-767-33178  
; Sequence 33178, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 33178  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-33178  
Alignment Scores:  
Pred. No.: 0.00445 Length: 151  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.80% Indels: 0  
DB: 4 Gaps: 0  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-09-270-767-33178 (1-151)  
Qy 451 GACAGTTGCCAGGGTGACAGTGGAGGTCCTCTGGTT 486  
|||||  
Db 72 AspSerCysGlnGlyAspSerGlyGlyProLeuVal 83  
RESULT 9  
US-09-270-767-48395  
; Sequence 48395, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 48395  
; LENGTH: 151  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
; FEATURE:  
; OTHER INFORMATION: Xaa means any amino acid  
US-09-270-767-48395  
Alignment Scores:  
Pred. No.: 0.00445 Length: 151  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.80% Indels: 0  
DB: 4 Gaps: 0  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-09-270-767-48395 (1-151)  
Qy 451 GACAGTTGCCAGGGTGACAGTGGAGGTCCTCTGGTT 486  
|||||  
Db 72 AspSerCysGlnGlyAspSerGlyGlyProLeuVal 83

RESULT 10  
US-09-518-046-24  
; Sequence 24, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; EARLIER FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 24  
; LENGTH: 159  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DOMAIN  
; OTHER INFORMATION: protease domain of TMPRSS2 (Tmprs2)  
US-09-518-046-24  
Alignment Scores:  
Pred. No.: 0.0042 Length: 159  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.80% Indels: 0  
DB: Gaps: 0  
US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x US-09-518-046-24 (1-159)  
QY 451 GACAGTTGCCAGGTGACAGTGGAGTCTCTCTGTT 486  
Db 147 AspSerCysGlnGlyAspSerGlyGlyProLeuVal 158  
RESULT 11  
US-09-879-792-16  
; Sequence 16, Application US/09879792  
; Patent No. 6734006  
; GENERAL INFORMATION:  
; APPLICANT: Xiaohong, Yonghong  
; APPLICANT: Gedrich, Richard  
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine  
; TITLE OF INVENTION: Protease  
; FILE REFERENCE: 02973.00035  
; CURRENT APPLICATION NUMBER: US/09/879,792  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/211,224  
; PRIOR FILING DATE: 2000-06-13  
; PRIOR APPLICATION NUMBER: US 60/283,353  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: US 60/283,648  
; PRIOR FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: PCT  
; PRIOR FILING DATE: 2001-06-12  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: BLOCKS BL1253G  
US-09-879-792-16  
Alignment Scores:  
Pred. No.: 0.0563 Length: 12  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: Gaps: 0  
US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x US-09-879-792-16 (1-12)  
QY 457 TGCCAGGTGACAGTGGAGTCTCTGTTGTC 489  
Db 2 CysGlnGlyAspSerGlyGlyProLeuValCys 12  
RESULT 12  
US-08-906-769-99  
; Sequence 99, Application US/08906769  
; Patent No. 6077687  
; GENERAL INFORMATION:  
; APPLICANT: Grieve, Robert B.  
; APPLICANT: Rushlow, Keith E.  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Frank, Glenn R.  
; APPLICANT: Stiegler, Gary  
; APPLICANT: Gaines, Patrick J.  
; APPLICANT: Silver, Gary  
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
; TITLE OF INVENTION: MOLECULES AND USES THEREOF  
; NUMBER OF SEQUENCES: 190  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/906,769  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/639,075  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-25-C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 99:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 84 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-906-769-99  
Alignment Scores:  
Pred. No.: 0.0433 Length: 84  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: Gaps: 0  
US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x US-08-906-769-99 (1-84)  
QY 451 GACAGTTGCCAGGTGACAGTGGAGTCTCTG 483  
Db 74 AspSerCysGlnGlyAspSerGlyGlyProLeu 84

```
RESULT 13
US-08-906-616-99
; Sequence 99, Application US/08906616
; Patent No. 6121035
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-906-616-99
Alignment Scores:
Pred. No.: 0.0433 Length: 84
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.31% Indels: 0
DB: 3 Gaps: 0
US-09-992-095B-53_COPY_1044_1664 (1-621) x US-08-906-616-99 (1-84)
Qy 451 GACAGTTCACGGTGACAGTGAGGTCCTCTG 483
Db 74 AspSerCysGlnGlyAspSerGlyGlyProLeu 84
RESULT 14
US-08-817-795-99
; Sequence 99, Application US/08817795
; Patent No. 6139840
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Frank, Glenn R.
; APPLICANT: Heath, Andrew W.
; APPLICANT: Yamaka, Miles Yamanaka
; APPLICANT: Arfsten, Ann
; APPLICANT: Dale, Beverly
; APPLICANT: Stiegler, Gary
; TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,795
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: PCT/US95/14442
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gary J. Connell
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-817-795-99
Alignment Scores:
Pred. No.: 0.0433 Length: 84
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.31% Indels: 0
DB: 3 Gaps: 0
US-09-992-095B-53_COPY_1044_1664 (1-621) x US-08-817-795-99 (1-84)
Qy 451 GACAGTTCACGGTGACAGTGAGGTCCTCTG 483
Db 74 AspSerCysGlnGlyAspSerGlyGlyProLeu 84
RESULT 15
US-08-639-075A-99
; Sequence 99, Application US/08639075A
; Patent No. 6150125
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
```



```

; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/639,075A
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-639-075A-99

```

```

Alignment Scores:
Pred. No.: 0.0433 Length: 84
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.31% Indels: 0
DB: 3 Gaps: 0

```

US-09-992-095b-53\_copy\_1044\_1664 (1-621) x US-08-639-075A-99 (1-84)

```

Qy 451 GACAGTTGCCAGGTGACAGTGAGGTCTCTG 483
Db 74 AppSerCysGlnGlyAspSerGlyGlyProLeu 84

```

Search completed: October 27, 2004, 09:13:22  
Job time : 24.5 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 27, 2004, 09:11:47 : Search time 208.5 Seconds  
(without alignments)  
1931.300 Million cell updates/sec

Title: US-09-992-095b-53\_COPY\_1044\_1664  
Perfect score: 207  
Sequence: 1 atgcactctctggaggcac.....agggagtgtatgagaataat 621

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1370721 seqs, 324215800 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1529548

Minimum DB seq length: 0  
Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Command line parameters:

-O=/cgn2\_1/USPTO.spool/US0992095/runat\_26102004\_085837\_10823/app\_query.fasta\_1.775  
-DB=Published Applications AA -QWMT=Eastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=207 -USER=US0992095\_@CGN\_1\_1\_86\_@runat\_26102004\_085837\_10823 -NCPU=6  
-ICPU=3 -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US12\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	207	100.0	207	10	US-09-992-600A-54	Sequence 54, Appl
2	207	100.0	207	10	US-09-924-340-54	Sequence 54, Appl
3	207	100.0	207	10	US-09-992-095B-54	Sequence 54, Appl
4	207	100.0	207	10	US-09-999-570-54	Sequence 54, Appl
5	207	100.0	207	14	US-10-000-489-54	Sequence 54, Appl
6	207	100.0	207	14	US-10-000-986-54	Sequence 54, Appl
7	207	100.0	207	14	US-10-154-678-54	Sequence 54, Appl
8	14	6.8	118	10	US-09-997-003-39	Sequence 39, Appl
9	13	6.3	18	9	US-09-826-290-247	Sequence 247, App
10	13	6.3	18	15	US-10-284-309-68	Sequence 68, Appl
11	13	6.3	23	9	US-09-879-792-22	Sequence 22, Appl
12	13	6.3	23	17	US-10-806-370-22	Sequence 22, Appl
13	13	6.3	23	9	US-09-826-290-130	Sequence 120, App
14	13	6.3	29	9	US-09-825-751A-8	Sequence 8, Appli
15	13	6.3	149	14	US-10-357-175-20	Sequence 20, Appl
16	13	6.3	149	14	US-10-455-720-20	Sequence 20, Appl
17	13	6.3	157	14	US-10-357-175-23	Sequence 23, Appl
18	13	6.3	157	14	US-10-455-720-23	Sequence 23, Appl
19	13	6.3	191	13	US-10-045-367A-6	Sequence 6, Appli
20	13	6.3	191	14	US-10-170-789-56	Sequence 56, Appl
21	13	6.3	193	14	US-10-106-698-4740	Sequence 4740, Ap
22	13	6.3	199	14	US-10-131-409-98	Sequence 98, Appl
23	13	6.3	199	14	US-10-139-854-98	Sequence 98, Appl
24	13	6.3	199	14	US-10-150-813-98	Sequence 98, Appl
25	13	6.3	199	15	US-10-150-811-98	Sequence 98, Appl
26	12	5.8	12	9	US-09-804-156-33	Sequence 33, Appl
27	12	5.8	12	9	US-09-946-633-18	Sequence 18, Appl
28	12	5.8	12	10	US-09-997-003-53	Sequence 53, Appl
29	12	5.8	12	13	US-10-125-459-18	Sequence 18, Appl
30	12	5.8	12	13	US-10-067-761-33	Sequence 33, Appl
31	12	5.8	12	14	US-10-319-519-33	Sequence 33, Appl
32	12	5.8	12	14	US-10-357-175-24	Sequence 24, Appl
33	12	5.8	159	14	US-10-455-720-24	Sequence 24, Appl
34	11	5.3	12	9	US-09-879-792-16	Sequence 16, Appl
35	11	5.3	12	17	US-10-806-370-16	Sequence 16, Appl
36	11	5.3	30	9	US-09-820-893-111	Sequence 111, App
37	11	5.3	30	15	US-10-607-565-111	Sequence 2, Appli
38	11	5.3	98	14	US-10-240-730-2	Sequence 2, Appli
39	11	5.3	102	15	US-10-262-511-56	Sequence 56, Appl
40	11	5.3	119	15	US-10-262-511-54	Sequence 54, Appl
41	11	5.3	119	15	US-10-262-511-60	Sequence 60, Appl
42	11	5.3	142	15	US-10-262-511-48	Sequence 48, Appl
43	11	5.3	148	15	US-10-262-511-52	Sequence 52, Appl
44	11	5.3	162	14	US-10-325-745-6	Sequence 6, Appli
45	11	5.3	187	9	US-09-804-156-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-09-992-600A-54  
; Sequence 54, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjanin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US4.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574

```

; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: JPatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-992-600A-54

Alignment Scores:
Pred. No.: 4,27e-192 Length: 207
Score: 207.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-09-992-600A-54 (1-207)

QY 1 ATGCACCTCTGGAGGACCTTGATATCCCGAGTGGGTGTGACTGCTGCCACTGC 60
DB 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20

QY 61 TTGGAGAAGTCCCAAGGCTTCATCTACAGGTCTCCTGGGTGCACACCAAGAAGTG 120
DB 21 LeuGluLysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40

QY 121 AATCTCGAACCCCATGTTTACGAAATAGAAAGTGTCTAGGCTGTTTTCAGAGCCACACGA 180
DB 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60

QY 181 AAGATATTGCTTCTTAAGTAGCAGTCTGCTCCCTCATCTGACAAAGTAATCCCA 240
DB 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80

QY 241 GCTTGTCTGCCATCCCAAAATATGTTGGTGTGACCGGACCGAATGTTTCATCACTGGC 300
DB 81 AlaCysLeuProSerProAsnTyrlsValValAlaAspArgThrGluCysPheIleThrGly 100

QY 301 TGGGAGAAACCCAGGTACTTTTGGAGTGGCTTCTCAAGAGGCCAGCTCCCTGTG 360
DB 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120

QY 361 ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGAGAGTCAATCCACCGAA 420
DB 121 IleGluAsnLysValCysAsnArgTyrlsPheLeuAsnGlyArgValGlnSerThrGlu 140

QY 421 CTCTGTCTGGGCATTTGGCCGAGGCACTGACAGTTCGAGGTGCAGTGGAGTCTCT 480
DB 141 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 160

QY 481 CTGGTTTGTCTCGAGAAGCAAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTGGC 540
DB 161 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180

QY 541 TGTGCACGCCCAATAAGCTGTGTCTATGTTTCTGTTTCAAGTTTGTACTTGGATT 600
DB 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200

QY 601 GAGGAGTGTGAGAAATAAT 621
DB 201 GluGlyValMetArgAsnAsn 207

```

RESULT 2

```

US-09-924-340-54
; Sequence 54, Application US/09924340
; Publication No. US20030027248A1
; GENERAL INFORMATION:
; APPLICANT: Bejanin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 91.US2.REG
; CURRENT APPLICATION NUMBER: US/09/924,340

```

```

; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-924-340-54

```

```

Alignment Scores:
Pred. No.: 4,27e-192 Length: 207
Score: 207.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-09-924-340-54 (1-207)

QY 1 ATGCACCTCTGGAGGACCTTGATATCCCGAGTGGGTGTGACTGCTGCCACTGC 60
DB 1 MethHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20

QY 61 TTGGAGAAGTCCCAAGGCTTCATCTACAGGTCTCCTGGGTGCACACCAAGAAGTG 120
DB 21 LeuGluLysSerProArgProSerSerTyrlsValIleLeuGlyAlaHisGlnGluVal 40

QY 121 AATCTCGAACCCCATGTTTACGAAATAGAAAGTGTCTAGGCTGTTTTCAGAGCCACACGA 180
DB 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60

QY 181 AAGATATTGCTTCTTAAGTAGCAGTCTGCTCCCTCATCTGACAAAGTAATCCCA 240
DB 61 LysAspIleAlaLeuLeuLysLeuSerProAlaValIleThrAspLysValIlePro 80

QY 241 GCTTGTCTGCCATCCCAAAATATGTTGGTGTGACCGGACCGAATGTTTCATCACTGGC 300
DB 81 AlaCysLeuProSerProAsnTyrlsValValAlaAspArgThrGluCysPheIleThrGly 100

QY 301 TGGGAGAAACCCAGGTACTTTTGGAGTGGCTTCTCAAGAGGCCAGCTCCCTGTG 360
DB 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120

QY 361 ATTGAGAATAAGTGTGCAATCGCTATGAGTTTCTGAATGAGAGTCAATCCACCGAA 420
DB 121 IleGluAsnLysValCysAsnArgTyrlsPheLeuAsnGlyArgValGlnSerThrGlu 140

QY 421 CTCTGTCTGGGCATTTGGCCGAGGCACTGACAGTTCGAGGTGCAGTGGAGTCTCT 480
DB 141 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 160

QY 481 CTGGTTTGTCTCGAGAAGCAAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTGGC 540
DB 161 LeuValCysPheGluLysAspLysTyrlsLeuGlnGlyValThrSerTrpGlyLeuGly 180

QY 541 TGTGCACGCCCAATAAGCTGTGTCTATGTTTCTGTTTCAAGTTTGTACTTGGATT 600
DB 181 CysAlaArgProAsnLysProGlyValTyrlsValArgValSerArgPheValThrTrpIle 200

QY 601 GAGGAGTGTGAGAAATAAT 621
DB 201 GluGlyValMetArgAsnAsn 207

RESULT 3
US-09-992-095B-54

```



Db 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120  
QY 361 ATTGAGATAAAGTGTGAATCGCTATGAGTTTCTGAATGAAGAGTCCCAATCCACCGAA 420  
Db 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140  
QY 421 CTCTGCTGGGCATTTCGCGGAGGCACCTGACAGTTCGCGGGTGACAGTGGAGGTCCT 480  
Db 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160  
QY 481 CTGGTTTGTCTTCGAGAGGACAAATACATTTTACAAGGAGTACCTTCTTGGGGTCTTGGC 540  
Db 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180  
QY 541 TGTGACGCCCAATAAGCCTGCTATGTTCTGTTTCAAGTTTGTACTTGGATT 600  
Db 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200  
QY 601 GAGGAGTGTATGAGAAATAAT 621  
Db 201 GluGlyValMetArgAsnAsn 207

## RESULT 5

US-10-000-489-54  
; Sequence 54, Application US/10000489  
; Publication No. US2003009201A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

; FILE REFERENCE: 91 US6 DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; PRIORITY FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-54

Alignment Scores:  
Pred. No.: 4,278-192 Length: 207  
Score: 207.00 Matches: 207  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x US-10-000-489-54 (1-207)

QY 1 ATGCACTTCTGTGAGGACCTTGATATCCCGAGAGTGGTGTGACTGCTGCCACTGC 60  
Db 1 MetHisPheCysGlyGlyThrLeuLysSerProGluTrpValLeuThrAlaAlaHisCys 20  
QY 61 TTGGAGAGTCCCAAGGCTTCACTTCACAGTTCCTCGGTGCGACACCAAGTGT 120  
Db 21 LeuGluLysSerProArgProSerTyrIysValIleLeuGlyAlaHisGlnGluVal 40  
QY 121 AATCTCGAAGCCGATGTTTCAGGAAATAGAGTGTCTAGCTGTTCTTGGAGCCACAGA 180

Db 41 AsnLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60  
QY 181 AAGATATTGCTTGTCTAAAGCTAAGCAGTCCCTCGCTCATCTGACAAAGTAAATCCCA 240  
Db 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80  
QY 241 GCTTGTCTGCGCATCCCAATTTATGTGCTGCTGACCGGACCGCAATGTTTCATCACTGGC 300  
Db 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100  
QY 301 TGGGAGAAAACCAAGGTACTTTTGGAGTGGCTTCTCAAGGAAGCCCAAGTCCCTGTG 360  
Db 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120  
QY 361 ATTGAGATAAAGTGTCTCAATCGCTATGAGTTTCTGAATGAAGAGTCCCAATCCACCGAA 420  
Db 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140  
QY 421 CTCTGCTGCTGGCATTTGGCGGAGGCACCTGACAGTTTGCAGGGTGACAGTGGAGGTCCT 480  
Db 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160  
QY 481 CTGTTTGTCTTCGAGAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGC 540  
Db 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180  
QY 541 TGTGACGCCCAATAAGCCTGCTATGTTCTGTTTCAAGTTTGTACTTGGATT 600  
Db 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200  
QY 601 GAGGAGTGTATGAGAAATAAT 621  
Db 201 GluGlyValMetArgAsnAsn 207

## RESULT 6

US-10-000-986-54  
; Sequence 54, Application US/10000986  
; Publication No. US20030096247A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91 US9 DIV  
; CURRENT APPLICATION NUMBER: US/10/000,986  
; PRIORITY FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 54  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-986-54

Alignment Scores:  
Pred. No.: 4,278-192 Length: 207  
Score: 207.00 Matches: 207  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-992-095b-53\_copy\_1044\_1664 (1-621) x US-10-000-986-54 (1-207)

```
QY 1 ATGCACCTTCTGGAGGACCTTGATATCCCGAGAGTGGTGTGACTGCTGCCACTGC 60
Db 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
QY 61 TTGGAGAAGTCCCAAGGCTTTCATCTCAAGGTCTCTGGGTGCACACCAAGAAGTG 120
Db 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
QY 121 AATCTCGAACCCGATCTTCAGGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 180
Db 41 AenLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
QY 181 AAGATATTGCCTTGTAAAGCTTAAGCAGTCCCTGCGCTCATCTCACAAAGTAATCCCA 240
Db 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
QY 241 GCTTGTCTGCATCCCAATATTATGTGGTCTGCTGACCGAGCCGAAATGTTTCATCACTGGC 300
Db 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
QY 301 TGGGAGAAACCAAGTACTTTTGGAGTGGCTTCTCAAGGAAGCCAGTCCCTGTG 360
Db 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
QY 361 ATTGAGAAATAGTGCATCTGCTATGAGTGTCTTCAAGGAGTCACTTCTGGGTCTTGGC 540
Db 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
QY 541 TGTGACGCCCAATAGCTGTCTATGTTCTGTTCAAGGTTTCAAGGTTTGTACTTGATT 600
Db 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
QY 601 GAGGGAGTGATGAGAAATAAT 621
Db 201 GluGlyValMetArgAsn 207
```

RESULT 7

```
US-10-154-678-54
; Sequence 54, Application US/10154678
; Publication No. US20030162186A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephane
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 182 US1 REG
; CURRENT APPLICATION NUMBER: US/10/154,678
; PRIORITY FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 09/924,340
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/305,456
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/302,277
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/298,698
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/293,574
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: JPatent
; SEQ ID NO 54
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-154-678-54

```
Alignment Scores:
Pred. No.: 4,27e-192 Length: 207
Score: 207.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-992-095b-53_copy_1044_1664 (1-621) x US-10-154-678-54 (1-207)
QY 1 ATGCACCTTCTGGAGGACCTTGATATCCCGAGAGTGGTGTGACTGCTGCCACTGC 60
Db 1 MetHisPheCysGlyThrLeuIleSerProGluTrpValLeuThrAlaAlaHisCys 20
QY 61 TTGGAGAAGTCCCAAGGCTTTCATCTCAAGGTCTCTGGGTGCACACCAAGAAGTG 120
Db 21 LeuGluLysSerProArgProSerTyrLysValIleLeuGlyAlaHisGlnGluVal 40
QY 121 AATCTCGAACCCGATCTTCAGGAAATAGAGTGTCTAGGCTGTCTTGGAGCCACACGA 180
Db 41 AenLeuGluProHisValGlnGluIleGluValSerArgLeuPheLeuGluProThrArg 60
QY 181 AAGATATTGCCTTGTAAAGCTTAAGCAGTCCCTGCGCTCATCTCACAAAGTAATCCCA 240
Db 61 LysAspIleAlaLeuLeuLysLeuSerSerProAlaValIleThrAspLysValIlePro 80
QY 241 GCTTGTCTGCATCCCAATATTATGTGGTCTGCTGACCGAGCCGAAATGTTTCATCACTGGC 300
Db 81 AlaCysLeuProSerProAsnTyrValValAlaAspArgThrGluCysPheIleThrGly 100
QY 301 TGGGAGAAACCAAGTACTTTTGGAGTGGCTTCTCAAGGAAGCCAGTCCCTGTG 360
Db 101 TrpGlyGluThrGlnGlyThrPheGlyAlaGlyLeuLeuLysGluAlaGlnLeuProVal 120
QY 361 ATTGAGAAATAGTGCATCTGCTATGAGTGTCTTCAAGGAGTCCATCCATCCACGAA 420
Db 121 IleGluAsnLysValCysAsnArgTyrGluPheLeuAsnGlyArgValGlnSerThrGlu 140
QY 421 CTCTGTCTGGGCAATTTGCCGGAGGACCTGACAGTTGCCAGGTGCACAGTGGAGTCCCT 480
Db 141 LeuCysAlaGlyHisLeuAlaGlyThrAspSerCysGlnGlyAspSerGlyGlyPro 160
QY 481 CTGTTGTCTGAGAGGACCAATAATACATTTTCAAGGAGTCACTTCTTGGGTCTTGGC 540
Db 161 LeuValCysPheGluLysAspLysTyrIleLeuGlnGlyValThrSerTrpGlyLeuGly 180
QY 541 TGTGACGCCCAATAGCTGTCTATGTTCTGTTCAAGGTTTCAAGGTTTGTACTTGATT 600
Db 181 CysAlaArgProAsnLysProGlyValTyrValArgValSerArgPheValThrTrpIle 200
QY 601 GAGGGAGTGATGAGAAATAAT 621
Db 201 GluGlyValMetArgAsn 207
```

RESULT 8

```
US-09-997-003-39
; Sequence 39, Application US/09997003
; Publication No. US20030203361A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA003P1
; CURRENT APPLICATION NUMBER: US/09/997,003
; PRIORITY FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/22157
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/148,680
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 56
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-003-39

Alignment Scores:
Pred. No.: 0.000189 Length: 118
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.76% Indels: 0
DB: 10 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-09-997-003-39 (1-118)

QY 448 ACTGACGTTCCAGGTCGACAGTGAGGTCCTCTGTTTC 489
DB 56 ThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 69

RESULT 9
US-09-826-290-247
; Sequence 247, Application US/09826290
; Patent No. US2002018468A1
; GENERAL INFORMATION:
; APPLICANT: Durham, L. Kathryn
; APPLICANT: Friedmann, David L.
; APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlff, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Therefor, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 247
; LENGTH: 18
; TYPE: PRT
; ORGANISM: homo sapien
US-09-826-290-247

Alignment Scores:
Pred. No.: 0.00239 Length: 18
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.28% Indels: 0
DB: 9 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-09-826-290-247 (1-18)

QY 451 GACAGTTCACAGGTCGACAGTGAGGTCCTCTGTTTC 489
DB 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13

RESULT 10
US-10-264-309-68
; Sequence 247, Application US/09826290
; Patent No. US2002018468A1
; GENERAL INFORMATION:
; APPLICANT: Durham, L. Kathryn
; APPLICANT: Friedmann, David L.
; APPLICANT: Herath, Herath Mudiyansele Athula Chandrasiri
; APPLICANT: Kimmel, Lida H.
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlff, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and Uses Therefor,
; TITLE OF INVENTION: INCLUDING DIAGNOSIS AND TREATMENT OF ALZHEIMER'S DISEASE
; FILE REFERENCE: POA-002.01
; CURRENT APPLICATION NUMBER: US/10/264,309
; CURRENT FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: 60/326,708
; PRIOR FILING DATE: 2001-10-03
; NUMBER OF SEQ ID NOS: 491
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 68
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-309-68

Alignment Scores:
Pred. No.: 0.00239 Length: 18
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.28% Indels: 0
DB: 15 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-10-264-309-68 (1-18)

QY 451 GACAGTTCACAGGTCGACAGTGAGGTCCTCTGTTTC 489
DB 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13

RESULT 11
US-09-879-792-22
; Sequence 22, Application US/09879792
; Patent No. US20020061850A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; APPLICANT: Gedrich, Richard
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; TITLE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/09/879,792
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
```



```

; FEATURE:
; OTHER INFORMATION: BLOCKS BL00134B
US-09-879-792-22

Alignment Scores:
Pred. No.: 0.0023 Length: 23
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.28% Indels: 0
DB: 9 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-09-879-792-22 (1-23)
QY 451 GACAGTTCAGGTCACAGTGGAGTCTCTGTTGC 489
DB 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13

RESULT 12
US-10-806-370-22
; Sequence 22, Application US/10806370
; Publication No. US20040209327A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human Transmembrane Serine
; TITLE OF INVENTION: Protease
; FILE REFERENCE: 02973.00035
; CURRENT APPLICATION NUMBER: US/10/806,370
; CURRENT FILING DATE: 2004-03-23
; PRIOR APPLICATION NUMBER: US/09/879,792
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/211,224
; PRIOR FILING DATE: 2000-06-13
; PRIOR APPLICATION NUMBER: US 60/283,353
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/283,648
; PRIOR FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: PCT
; PRIOR FILING DATE: 2001-06-12 (Docket No. LJO-81-WO)
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: BLOCKS BL00134B
US-10-806-370-22

Alignment Scores:
Pred. No.: 0.0023 Length: 23
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.28% Indels: 0
DB: 17 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-10-806-370-22 (1-23)
QY 451 GACAGTTCAGGTCACAGTGGAGTCTCTGTTGC 489
DB 1 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 13

RESULT 13
US-09-826-290-120
; Sequence 120, Application US/09826290
; Patent No. US20020164668A1
; GENERAL INFORMATION:
; APPLICANT: Durham, L. Kathryn
; APPLICANT: Friedman, David L.
; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
; APPLICANT: Kimmel, Lida H.

```

```

; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Potter, David M.
; APPLICANT: Rohlff, Christian
; APPLICANT: Silber, B. Michael
; APPLICANT: Stiger, Thomas R.
; APPLICANT: Sunderland, P. Trey
; APPLICANT: Townsend, Robert Reid
; APPLICANT: White, Frost
; APPLICANT: Williams, Stephen A.
; TITLE OF INVENTION: Nucleic Acid Molecules, Polypeptides and
; TITLE OF INVENTION: Uses Thereof, Including Diagnosis and Treatment of
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 2572-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/826,290
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/194,504
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: US 60/253,647
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 492
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 29
; TYPE: PRT
; ORGANISM: homo sapien
US-09-826-290-120

Alignment Scores:
Pred. No.: 0.00221 Length: 29
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.28% Indels: 0
DB: 9 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x US-09-826-290-120 (1-29)
QY 451 GACAGTTCAGGTCACAGTGGAGTCTCTGTTGC 489
DB 12 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 24

RESULT 14
US-09-825-751A-8
; Sequence 8, Application US/09825751A
; Publication No. US20030065140A1
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Taupier, Raymond J.
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Rastelli, Luca
; APPLICANT: Herrman, John L.
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-750
; CURRENT APPLICATION NUMBER: US/09/825,751A
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/194,314
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 60/225,693
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-825-751A-8

Alignment Scores:
Pred. No.: 0.00193 Length: 97
Score: 13.00 Matches: 13

```

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.28% Indels: 0  
 DB: 10 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-09-825-751A-8 (1-97)

Qy 451 GACAGTTCACAGGTGACAGTGGAGTCTCTGTTTGC 489  
 Db 41 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 53

RESULT 15

US-10-357-175-20  
 ; Sequence 20, Application US/10357175  
 ; Publication No. US20030170707A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Brien, Timothy J.  
 ; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
 ; FILE REFERENCE: D6192CIP/D/CIP  
 ; CURRENT APPLICATION NUMBER: US/10/357,175  
 ; CURRENT FILING DATE: 2003-02-03  
 ; PRIOR APPLICATION NUMBER: 09/650,371  
 ; PRIOR FILING DATE: 2000-08-28  
 ; NUMBER OF SEQ ID NOS: 158  
 ; SEQ ID NO 20  
 ; LENGTH: 149  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: DOMAIN  
 ; OTHER INFORMATION: protease domain of protease M (Prom)  
 US-10-357-175-20

Alignment Scores:  
 Pred. No.: 0.00171 Length: 149  
 Score: 13.00 Matches: 13  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.28% Indels: 0  
 DB: 14 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x US-10-357-175-20 (1-149)

Qy 451 GACAGTTCACAGGTGACAGTGGAGTCTCTGTTTGC 489  
 Db 137 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 149

Search completed: October 27, 2004, 09:27:25  
 Job time : 212.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: October 27, 2004, 01:40:30 ; Search time 3041 Seconds  
(without alignments)  
9656.989 Million cell updates/sec

Title: US-09-992-095B-53\_COPY\_1044\_1664  
Perfect score: 621  
Sequence: 1 atgcactctctggaggcac.....aggagtgatgagaaataat 621

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 4526729 seqs, 23644849745 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 5954484

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	149	24.0	254	9 HUMPLG23	M34275 Human plasm
2	141	22.7	331	9 HUMPLG20	M34272 Human plasm
3	129	20.8	356	6 CQ433162	CQ433162 Sequence
4	129	20.8	436	6 CQ506843	CQ506843 Sequence
5	129	20.8	440	6 CQ476920	CQ476920 Sequence
6	129	20.8	470	6 CQ425554	CQ425554 Sequence
7	116	18.7	468	6 CQ415403	CQ415403 Sequence
8	112	18.0	251	9 HUMPLG21	M34273 Human plasm
9	83	13.4	312	6 AR418027	AR418027 Sequence
10	83	13.4	312	6 AX978721	AX978721 Sequence
11	83	13.4	312	6 BD113580	BD113580 EST and e
12	70	11.3	361	9 HUMPLG19	M33290 Human plasm
13	60	9.7	60	6 AX611981	AX611981 Sequence
14	60	9.7	60	6 AX611983	AX611983 Sequence
15	53	8.5	567	4 BOVPMG	K02935 Bovine plas
16	48	7.7	48	6 AX611977	AX611977 Sequence
17	48	7.7	48	6 AX611979	AX611979 Sequence
18	44	7.1	44	6 AX611985	AX611985 Sequence
19	41	6.6	586	11 BV100086	BV100086 RPAMWSEQO

20	41	6.6	592	11	BV159185
21	35	5.6	35	6	AX07993
22	35	5.6	60	6	AX611982
23	35	5.6	468	6	CQ434098
24	35	5.6	523	6	CQ725849
25	34	5.5	34	6	AX07992
26	32	5.2	40	6	AX611984
27	30	4.8	48	6	AX611980
28	29	4.7	476	6	CQ424108
29	26	4.2	299	6	AX312474
30	24	3.9	44	6	AX611986
31	24	3.9	48	6	AX611978
32	23	3.7	219	10	MMTSP1G2
33	23	3.7	540	10	MUSGRAA02
34	23	3.7	549	11	BV099297
35	23	3.7	549	11	BV159513
36	21	3.4	38	6	AX463629
37	20	3.2	20	11	DOGPLGA
38	20	3.2	20	11	DOGPLGB
39	20	3.2	39	6	I34472
40	20	3.2	39	6	I34474
41	20	3.2	252	6	CQ439589
42	20	3.2	316	11	AU027683
43	20	3.2	318	11	AU028271
44	19	3.1	33	6	AX815928
45	19	3.1	36	6	I34476

ALIGNMENTS

RESULT 1	HUMPLG23	254 bp	DNA	linear	PRI 07-JAN-1995
LOCUS	Human plasminogen gene, exon 18.				
DEFINITION	M34275 J05286				
ACCESSION	M34275.1 GI:190063				
VERSION	plasm1nogen.				
KEYWORDS	23 of 24				
SEGMENT					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 254)				
AUTHORS	Petersen, T.E., Martzen, M.R., Ichinose, A. and Davie, E.W.				
TITLE	Characterization of the gene for human plasminogen, a key proenzyme in the fibrinolytic system				
JOURNAL	J. Biol. Chem. 265 (11), 6104-6111 (1990)				
MEDLINE	90202879				
PUBMED	2318848				
COMMENT	Original source text: Homo sapiens DNA. Draft entry and computer-readable sequence for [1] kindly submitted by A. Ichinose, 26-MAR-1990, for release after publication.				

Original source text: Homo sapiens DNA.  
Draft entry and computer-readable sequence for [1] kindly submitted by A. Ichinose, 26-MAR-1990, for release after publication.

FEATURES	Location/Qualifiers
source	1..254
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/map="eq26-q27"
	/cell_type="leukocyte and lung fibroblast"
	<1..68
intron	
	/genes="PLG"
	/note="G00-119-498"
	/number=17
	69..214
exon	
	/genes="PLG"
	/note="G00-119-498"
	/number=18
	215..>254
intron	
	/genes="PLG"
	/note="G00-119-498"
	/number=18
ORIGIN	Undetermined number of base pairs after segment 22.

```
Query Match      24.0%; Score 149; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 7e-72;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 315 AGGTACTTTGGAGCTGGCCCTCTCAAGGAAGCCAGCTCCCTGTGATTGAGATAAAGT 374
Db 67 AGGTACTTTGGAGCTGGCCCTCTCAAGGAAGCCAGCTCCCTGTGATTGAGATAAAGT 126

Qy 375 GTGCAATCGCTATGAGTTCTTGAATGGAAGAGTCCAATCCACCGAACTCTGTGTGGGCA 434
Db 127 GTGCAATCGCTATGAGTTCTTGAATGGAAGAGTCCAATCCACCGAACTCTGTGTGGGCA 186

Qy 435 TTGGCCGAGGACATGACAGTTGCCAGG 463
Db 187 TTGGCCGAGGACATGACAGTTGCCAGG 215

RESULT 2
HUMPLG20      331 bp      DNA      linear      PRI 07-JAN-1995
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
Patent: WO 0151628-A 18196 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
source
1.356
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      20.8%; Score 129; DB 6; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.2e-60;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 493 GAGAAGCACAATAACATTTTACAAGAGTCACTTCTTGGGGTCTTGGCTGTCACGCCCC 552
Db 182 GAGAAGCACAATAACATTTTACAAGAGTCACTTCTTGGGGTCTTGGCTGTCACGCCCC 241

Qy 553 AATAAGCTGGTGTCTATGTTGTTTCAGGTTTCTTACTTGGATTGAGGAGTGTATG 612
Db 242 AATAAGCTGGTGTCTATGTTGTTTCAGGTTTCTTACTTGGATTGAGGAGTGTATG 301

Qy 613 AGAAATAAT 621
Db 302 AGAAATAAT 310

RESULT 4
CQ506843      436 bp      DNA      linear      PAT 30-JAN-2004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Schlegel, R., Endege, W.O. and Monahan, J.E.
Genes differentially expressed in human prostate cancer and their
use
Patent: WO 0160860-A 38710 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
Location/Qualifiers
source
1.436
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      20.8%; Score 129; DB 6; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.2e-60;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 GTCCCCAAGCCCTTCACTACAGTCTCTGGTGACACACCGAAGTCAATCTCGA 128
Db 41 GTCCCCAAGCCCTTCACTACAGTCTCTGGTGACACACCGAAGTCAATCTCGA 100

Qy 129 ACCGATGTTTCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGAAAGATAT 188
Db 101 ACCGATGTTTCAGGAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGAAAGATAT 160
```

QY 493 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 552  
| | | | |  
Db 229 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 288  
| | | | |  
QY 553 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 612  
| | | | |  
Db 289 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 348  
| | | | |  
QY 613 AGAATAAT 621  
| | | | |  
Db 349 AGAATAAT 357  
| | | | |  
RESULT 5  
LOCUS CQ476920 440 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 8787 from Patent WO0160860.  
ACCESSION CQ476920  
VERSION CQ476920.1 GI:41442539  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.  
TITLE Genes differentially expressed in human prostate cancer and their use  
JOURNAL Patent: WO 0160860-A 8787 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1. .440  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 20.8%; Score 129; DB 6; Length 440;  
Best Local Similarity 100.0%; Pred. No. 1.2e-60;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 493 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 552  
| | | | |  
Db 228 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 287  
| | | | |  
QY 553 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 612  
| | | | |  
Db 288 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 347  
| | | | |  
QY 613 AGAATAAT 621  
| | | | |  
Db 348 AGAATAAT 356  
| | | | |  
RESULT 6  
LOCUS CQ425554 470 bp DNA linear PAT 28-JAN-2004  
DEFINITION Sequence 10588 from Patent WO0151628.  
ACCESSION CQ425554  
VERSION CQ425554.1 GI:41377783  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.  
TITLE Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer  
JOURNAL Patent: WO 0151628-A 10588 19-JUL-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source Location/Qualifiers

1. .470  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
source  
ORIGIN  
Query Match 20.8%; Score 129; DB 6; Length 470;  
Best Local Similarity 100.0%; Pred. No. 1.2e-60;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 493 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 552  
| | | | |  
Db 186 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 245  
| | | | |  
QY 553 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 612  
| | | | |  
Db 246 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 305  
| | | | |  
QY 613 AGAATAAT 621  
| | | | |  
Db 306 AGAATAAT 314  
| | | | |  
RESULT 7  
LOCUS CQ415403 468 bp DNA linear PAT 28-JAN-2004  
DEFINITION Sequence 437 from Patent WO0151628.  
ACCESSION CQ415403  
VERSION CQ415403.1 GI:41367632  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.  
TITLE Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer  
JOURNAL Patent: WO 0151628-A 437 19-JUL-2001;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
source Location/Qualifiers  
1. .468  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 18.7%; Score 116; DB 6; Length 468;  
Best Local Similarity 100.0%; Pred. No. 2.5e-53;  
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 493 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 552  
| | | | |  
Db 201 GAGAGGACAAATACATTTTACAGGAGTCACCTTCTTGGGCTCTTGGCTGTGCACGCCCC 260  
| | | | |  
QY 553 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 608  
| | | | |  
Db 261 AATAAGCCTGGTCTATGTTCTGTTTCAAGGTTTGTCTTGGATTGAGGAGTGATG 316  
| | | | |  
RESULT 8  
LOCUS HUMPLG21 251 bp DNA linear PRI 07-JAN-1995  
DEFINITION Human plasminogen gene, exon 17.  
ACCESSION M34273 J05286  
VERSION M34273.1 GI:190061  
KEYWORDS plasminogen.  
SEGMENT 21 of 24  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 251)

AUTHORS Petersen,T.E., Martzen,M.R., Ichinose,A. and Davie,E.W.  
TITLE Characterization of the gene for human plasminogen, a key proenzyme in the fibrinolytic system  
JOURNAL J. Biol. Chem. 265 (11), 6104-6111 (1990)  
MEDLINE 90202879  
PUBMED 2318848  
COMMENT Original source text: Homo sapiens DNA.  
Draft entry and computer-readable sequence for [1] kindly submitted by A.Ichinose, 26-MAR-1990, for release after publication.

FEATURES  
Location/Qualifiers  
source  
1..251  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="6q26-q27"  
/cell\_type="leukocyte and lung fibroblast"  
intron  
1..40  
/gene="PLG"  
/note="G00-119-498"  
/number=16  
41..147  
/gene="PLG"  
/note="G00-119-498"  
/number=17  
148..>251  
/gene="PLG"  
/note="G00-119-498"  
/number=17  
exon  
148..>251  
intron  
148..>251  
ORIGIN  
Undetermined number of base pairs after segment 20.

Query Match 18.0%; Score 112; DB 9; Length 251;  
Best Local Similarity 100.0%; Pred. No. 4.8e-51;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 207 CAGTCTGCGTCATCACTGACAAAGTAATCCACGTTGTCTGCCATCCCAAAATTATGT 266  
DB 38 CAGTCTGCGTCATCACTGACAAAGTAATCCACGTTGTCTGCCATCCCAAAATTATGT 97  
QY 267 GGTGCTGACCGGACGAATGTTTCATCACTGGCTGGGAGAAACCCCAAGT 318  
DB 98 GGTGCTGACCGGACGAATGTTTCATCACTGGCTGGGAGAAACCCCAAGT 149

RESULT 9  
AR418027  
LOCUS 312 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 9524 from patent US 6639063.  
ACCESSION AR418027  
VERSION AR418027.1 GI:40173137  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 312)  
AUTHORS Edwards,J.-B.D.M., Jobert,S. and Giordano,J.-Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 9524 28-OCT-2003;  
FEATURES Location/Qualifiers  
source  
1..312  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 13.4%; Score 83; DB 6; Length 312;  
Best Local Similarity 100.0%; Pred. No. 1e-34;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 363 TGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCCAATCCACCGAACT 422  
DB 1 TGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCCAATCCACCGAACT 60  
QY 423 CTGTGCTGGCATTGGCCGGAG 445

Db 61 CTGTGCTGGCATTGGCCGGAG 83  
RESULT 10  
AX978721  
LOCUS 312 bp DNA linear PAT 15-JAN-2004  
DEFINITION Sequence 9524 from Patent EP1104808.  
ACCESSION AX978721  
VERSION AX978721.1 GI:40984861  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Dumas Milne Edwards,J.B., Jobert,S. and Giordano,J.Y.  
TITLE ESTs and encoded human proteins  
JOURNAL Patent: EP 1104808-A 9524 06-JUN-2001;  
Genset (FR)  
FEATURES Location/Qualifiers  
source  
1..312  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 13.4%; Score 83; DB 6; Length 312;  
Best Local Similarity 100.0%; Pred. No. 1e-34;  
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 363 TGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCCAATCCACCGAACT 422  
DB 1 TGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGTCCCAATCCACCGAACT 60  
QY 423 CTGTGCTGGCATTGGCCGGAG 445  
DB 61 CTGTGCTGGCATTGGCCGGAG 83

RESULT 11  
BD113580  
LOCUS 312 bp DNA linear PAT 18-SEP-2002  
DEFINITION EST and encoded human protein.  
ACCESSION BD113580  
VERSION BD113580.1 GI:23208398  
KEYWORDS JP 2002010789-A/5657.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 312)  
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 5657 15-JAN-2002;  
COMMENT GENSET CORP  
OS Homo sapiens (human)  
PN JP 2002010789-A/5657  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 2000280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI  
GIORDANO  
PC C12N15/09,C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02,C12P21/08,C12Q1/68,C12N15/00,C12N5/00, PC  
C12N15/00  
CC n-a, g, c or t  
FH Key Location/Qualifiers  
FT misc\_feature 189.  
Location/Qualifiers  
1..312  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

```

ORIGIN
/db_xref="taxon:9606"

Query Match      13.4%; Score 83; DB 6; Length 312;
Best Local Similarity 100.0%; Pred. No. 1e-34;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 363 TGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAACT 422
Db 1 TGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAGAGTCCAATCCACCGAACT 60

QY 423 CTGTGCTGGGCATTTGGCCGGAG 445
Db 61 CTGTGCTGGGCATTTGGCCGGAG 83

RESULT 12
HUMPLG19          361 bp      DNA      linear      PRI 07-JAN-1995
LOCUS             Human plasminogen gene, exon 15.
ACCESSION         M33290 J05286
VERSION           M33290.1 GI:190059
KEYWORDS          plasminogen.
SEGMENT           19 of 24
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE            Petersen,T.E., Martzen,M.R., Ichinose,A. and Davie,E.W.
                  Characterization of the gene for human plasminogen, a key proenzyme
                  in the fibrinolytic system
JOURNAL           J. Biol. Chem. 265 (11), 6104-6111 (1990)
MEDLINE           90202879
PUBMED            2318948
COMMENT           Original source text: Homo sapiens DNA.
                  Draft entry and computer-readable sequence for [1] kindly submitted
                  by A.Ichinose, 26-MAR-1990, for release after publication.

FEATURES
source            1..361
                  /organism="Homo sapiens"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:9606"
                  /map="6q26-q27"
                  /cell_type="leukocyte and lung fibroblast"
                  <1..65
                  /genes="PLG"
                  /notes="G00-119-498"
                  /number=14
                  66..140
                  /genes="PLG"
                  /notes="G00-119-498"
                  /number=15
                  141..>361
                  /genes="PLG"
                  /notes="G00-119-498"
                  /number=15

intron            Undetermined number of base pairs after segment 18.

exon              Query Match      11.3%; Score 70; DB 9; Length 361;
                  Best Local Similarity 100.0%; Pred. No. 2.1e-27;
                  Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

intron            QY 1 ATGCACCTTCGTGGAGCACCTTGATATCCCGAGTGGGTGTGACTGTGCCCACTGC 60
                  Db 73 ATGCACCTTCGTGGAGCACCTTGATATCCCGAGTGGGTGTGACTGTGCCCACTGC 132
                  QY 61 TTGGAGAACT 70
                  Db 133 TTGGAGAACT 142

RESULT 13

```

```

AX6111981
LOCUS             Sequence 3006 from Patent WO02072882.
DEFINITION        AX6111981
ACCESSION         AX6111981
VERSION           AX6111981.1 GI:28407410
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE            Cullen,P. and Seedorf,U.
JOURNAL           Coronary chip
                  Patent: WO 02072882-A 3006 19-SEP-2002;
                  OGHAM GmbH (DE)
FEATURES           Location/Qualifiers
source            1..60
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"

ORIGIN
Query Match      9.7%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 CTGCCATCCCAAAATTATGTGCTGCTGACCGACCGAATGTTTCATCACTGCTGGGA 306
Db 1 CTGCCATCCCAAAATTATGTGCTGCTGACCGACCGAATGTTTCATCACTGCTGGGA 60

RESULT 14
AX6111983
LOCUS             Sequence 3008 from Patent WO02072882.
DEFINITION        AX6111983
ACCESSION         AX6111983
VERSION           AX6111983.1 GI:28407412
KEYWORDS          Homo sapiens (human)
SOURCE            Homo sapiens
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE            Cullen,P. and Seedorf,U.
JOURNAL           Coronary chip
                  Patent: WO 02072882-A 3008 19-SEP-2002;
                  OGHAM GmbH (DE)
FEATURES           Location/Qualifiers
source            1..60
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"

ORIGIN
Query Match      9.7%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 CTGCCATCCCAAAATTATGTGCTGCTGACCGACCGAATGTTTCATCACTGCTGGGA 306
Db 1 CTGCCATCCCAAAATTATGTGCTGCTGACCGACCGAATGTTTCATCACTGCTGGGA 60

RESULT 15
BOVPMG           Bovine plasminogen mRNA, 3' end.
LOCUS             BOVPMG
DEFINITION        BOVPMG
ACCESSION         K02935
VERSION           K02935.1 GI:163551
KEYWORDS          glycoprotein; plasminogen.
SOURCE            Bos taurus (cow)
ORGANISM          Bos taurus
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovinae; Bos.  
REFERENCE 1 (bases 1 to 567)  
AUTHORS Malinowski,D.P., Sadler,J.E. and Davie,E.W.  
TITLE Characterization of a complementary deoxyribonucleic acid coding  
for human and bovine plasminogen  
JOURNAL Biochemistry 23 (18), 4243-4250 (1984)  
MEDLINE 85023311  
PUBMED 6148961  
COMMENT Original source text: Bovine liver, cDNA to mRNA, clone 5.  
FEATURES  
    source  
        1..567  
            /organism="Bos taurus"  
            /mol\_type="mRNA"  
            /db\_xref="taxon:9913"  
            <1..567  
            /product="pmg mRNA"  
            <1..324  
            /notes="plasminogen"  
            /codon\_start=1  
            /protein\_id="AAA30714.1"  
            /db\_xref="GI:163552"  
            /translation="WGTEQGTGFGRLKEAHLPIENKVCNNEYLDGRVKPRELCAG  
            HLIGTSCQDSDGGLVCFEKDKYILQGVTSWGLGCARPNKPGVYRVSPYVPWIEE  
            TMRN"  
ORIGIN 128 bp upstream of HpaII site.  
Query Match 8.5%; Score 53; DB 4; Length 567;  
Best Local Similarity 100.0%; Pred. No. 7.3e-18;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 514 CAAGGAGTCACTTCTTGGGGTCTTGGGTGTGACGCCCAATAAGCCTGGTGT 566  
    |||||  
Db 214 CAAGGAGTCACTTCTTGGGGTCTTGGGTGTGACGCCCAATAAGCCTGGTGT 266  
    |||||  
Search completed: October 27, 2004, 06:51:03  
Job time : 3046 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 01:24:40 ; Search time 457 Seconds  
(without alignments)  
7133.240 Million cell updates/sec

Title: US-09-992-095b-53\_COPY\_1044\_1664  
Perfect score: 621  
Sequence: 1 atgcactctctgaggagcac.....aggagtgatgagaataat 621

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6423072

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Database : N Geneseq\_23Sep04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	20.8	356	4	AAL25726 Human bre
2	129	20.8	436	5	ABV38691 Human pro
3	129	20.8	440	5	ABV08796 Human pro
4	129	20.8	470	4	AAL18121 Human bre
5	129	20.8	508	9	ACH47938 Human lun
6	116	18.7	468	4	AAL07980 Human bre
7	107	17.2	363	9	ACH42916 Human bre
8	103	16.6	357	9	ACH42960 Human foe
9	35	5.6	468	4	AAL26661 Human bre
10	33	5.3	346	8	ABX47578 Bovine ES
11	30	4.8	384	9	ACH20439 Human adu
12	29	4.7	476	4	AAL16677 Human bre
13	27	4.3	45	3	AAA48440 Human mic
14	27	4.3	394	8	ABX45923 Bovine ES
15	27	4.3	468	8	ABX45686 Bovine ES
16	26	4.2	299	6	ABN77783 Human cat
17	25	4.0	162	8	ABX41917 Bovine ES
18	25	4.0	203	8	ABX46133 Bovine ES
19	25	4.0	234	8	ABX49218 Bovine ES
20	21	3.4	21	2	AAX78885 Human pla
21	21	3.4	38	6	ABN89463 Human mic

C	22	21	3.4	45	3	AAA48439	Aaa48439 Human mic
	23	21	3.4	384	8	ABX45828	Abx45828 Bovine ES
	24	20	3.2	20	2	AAV01185	Aav01185 Plasminog
C	25	20	3.2	20	2	AAV01186	Aav01186 Plasminog
	26	20	3.2	39	2	AAQ91617	Aaq91617 Human apo
	27	20	3.2	39	2	AAQ91615	Aaq91615 Human apo
C	28	20	3.2	41	12	ADL64239	Adl64239 Human ssn
	29	20	3.2	252	6	ABN18436	Abn18436 Human ORF
C	30	19	3.1	33	10	ACH00746	Ach00746 Plasminog
	31	19	3.1	36	2	AAQ91619	Aaq91619 Human apo
	32	19	3.1	39	2	AAQ91614	Aaq91614 Human apo
C	33	19	3.1	39	2	AAQ91612	Aaq91612 Human apo
	34	19	3.1	46	12	ADO71586	Ado71586 Primer fo
	35	19	3.1	165	10	ADC27716	Adc27716 Human col
	36	19	3.1	300	2	AAZ14255	Aaz14255 Human gen
C	37	18	2.9	18	12	ADO71590	Ado71590 Primer fo
	38	18	2.9	30	12	ADO71580	Ado71580 Primer fo
	39	18	2.9	39	2	AAQ91616	Aaq91616 Human apo
C	40	18	2.9	51	12	ADO71585	Ado71585 Primer fo
	41	18	2.9	51	12	ADO71584	Ado71584 Primer fo
	42	18	2.9	165	12	ACH81929	Ach81929 Human gen
	43	18	2.9	221	12	ACH87813	Ach87813 Human gen
C	44	18	2.9	232	8	ABT23431	Abt23431 Immune-re
C	45	18	2.9	300	2	AAZ13138	Aaz13138 Human gen

ALIGNMENTS

RESULT 1  
AAL25726  
ID AAL25726 standard; cdna; 356 BP.  
XX  
AC AAL25726;  
XX  
DT 07-DEC-2001 (first entry)  
XX  
DE Human breast cancer expressed polynucleotide 18183.  
XX  
KW Human; breast cancer; cell marker; cytostatic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200151628-A2.  
XX  
PD 19-JUL-2001.  
XX  
PF 10-JAN-2001; 2001WO-US000798.  
XX  
PR 14-JAN-2000; 2000US-0176077P.  
PR 14-MAR-2000; 2000US-0189167P.  
PR 24-MAR-2000; 2000US-0192099P.  
PR 29-MAR-2000; 2000US-0193480P.  
PR 15-MAY-2000; 2000US-0205230P.  
PR 09-JUN-2000; 2000US-0211315P.  
PR 25-JUL-2000; 2000US-0220534P.  
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
Lillie J, Xu Y, Wang Y, Steinmann K;  
WPI; 2001-451856/48.  
New peptide useful as a marker for the diagnosis of breast cancer.

Claim 1; Page 3360; 3695pp; English.  
The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and

CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity  
 XX  
 SQ Sequence 356 BP; 98 A; 71 C; 86 G; 101 T; 0 U; 0 Other;  
 Query Match 20.8%; Score 129; DB 4; Length 356;  
 Best Local Similarity 100.0%; Pred. No. 3e-57;  
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 493 GAGAGGACAAATACATTTTACAAGGAGTCACTTCTGGGGTCTTGGCTGTGCACGCC 552  
 DB 182 GGAAGGACAAATACATTTTACAAGGAGTCACTTCTGGGGTCTTGGCTGTGCACGCC 241  
 QY 553 AATAAGCCTGGTGTCTATGTTCTTCAAGGTTTGTACTTGGATTGAGGGAGTGATG 612  
 DB 242 AATAAGCCTGGTGTCTATGTTCTTCAAGGTTTGTACTTGGATTGAGGGAGTGATG 301  
 QY 613 AGAATAAT 621  
 DB 302 AGAATAAT 310

RESULT 2  
 ABV38691  
 ID ABV38691 standard; cDNA; 436 BP.  
 XX  
 AC ABV38691;  
 XX  
 DT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 38682.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 WPI; 2001-662795/76.  
 XX  
 DR Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 7870; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 436 BP; 117 A; 87 C; 116 G; 116 T; 0 U; 0 Other;  
 Query Match 20.8%; Score 129; DB 5; Length 436;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-57;  
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 493 GAGAGGACAAATACATTTTACAAGGAGTCACTTCTGGGGTCTTGGCTGTGCACGCC 552  
 DB 229 GAGAGGACAAATACATTTTACAAGGAGTCACTTCTGGGGTCTTGGCTGTGCACGCC 288  
 QY 553 AATAAGCCTGGTGTCTATGTTCTTCAAGGTTTGTACTTGGATTGAGGGAGTGATG 612  
 DB 289 AATAAGCCTGGTGTCTATGTTCTTCAAGGTTTGTACTTGGATTGAGGGAGTGATG 348  
 QY 613 AGAATAAT 621  
 DB 349 AGAATAAT 357

RESULT 3  
 ABV08796  
 ID ABV08796 standard; cDNA; 440 BP.  
 XX  
 AC ABV08796;  
 XX  
 DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 8787.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 WPI; 2001-662795/76.  
 XX  
 DR Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 1386; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX Sequence 440 BP; 115 A; 88 C; 120 G; 116 T; 0 U; 1 Other;  
SQ

Query Match 20.8%; Score 129; DB 5; Length 440;  
Best Local Similarity 100.0%; Pred. No. 3.1e-57;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 493 GAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCC 552  
DB 228 GAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCC 287  
QY 553 AATAAGCCTGGTGTCTATGTTTCAAGGTTTGTCTTCAAGTTTGTACTTGGATTGAGGAGTGATG 612  
DB 288 AATAAGCCTGGTGTCTATGTTTCAAGGTTTGTCTTCAAGTTTGTACTTGGATTGAGGAGTGATG 347  
QY 613 AGAATAAAT 621  
DB 348 AGAATAAAT 356

RESULT 4  
AAL18121  
ID AAL18121 standard; cDNA; 470 BP.  
AC AAL18121;  
XX  
DT 07-DEC-2001 (first entry)  
XX  
DE Human breast cancer expressed polynucleotide 10578.  
XX  
KW Human; breast cancer; cell marker; cytostatic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200151628-A2.  
XX  
PD 19-JUL-2001.  
XX  
PF 10-JAN-2001; 2001WO-US000798.  
XX  
PR 14-JAN-2000; 2000US-0176077P.  
PR 24-MAR-2000; 2000US-0189167P.  
PR 24-MAR-2000; 2000US-0192099P.  
PR 29-MAR-2000; 2000US-0193480P.  
PR 15-MAY-2000; 2000US-0205230P.  
PR 09-JUN-2000; 2000US-0211315P.  
PR 25-JUL-2000; 2000US-0220534P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Lillie J, Xu Y, Wang Y, Steinmann K;  
XX  
PS WPI; 2001-451856/48.  
XX  
PT New peptide useful as a marker for the diagnosis of breast cancer.  
XX  
PS Claim 1; Page 1890; 3695pp; English.  
XX

The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

Query Match 20.8%; Score 129; DB 4; Length 470;  
Best Local Similarity 100.0%; Pred. No. 3.1e-57;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 493 GAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCC 552  
DB 186 GAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTTGGCTGTGCACGCCCC 245  
QY 553 AATAAGCCTGGTGTCTATGTTTCAAGGTTTGTCTTCAAGTTTGTACTTGGATTGAGGAGTGATG 612  
DB 246 AATAAGCCTGGTGTCTATGTTTCAAGGTTTGTCTTCAAGTTTGTACTTGGATTGAGGAGTGATG 305  
QY 613 AGAATAAAT 621  
DB 306 AGAATAAAT 314

RESULT 5  
ACH47938/c  
ID ACH47938 standard; cDNA; 508 BP.  
XX  
AC ACH47938;  
XX  
DT 13-OCT-2003 (first entry)  
XX  
DE Human lung tumour cDNA #71.  
XX  
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST; genome mapping; biodiversity; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN US2003073623-A1.  
XX  
PD 17-APR-2003.  
XX  
PF 30-JUL-2001; 2001US-00918995.  
XX  
PR 30-JUL-2001; 2001US-00918995.  
XX  
PA (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
XX  
PS WPI; 2003-615964/58.  
XX

New polynucleotide sequences obtained from various cDNA libraries, useful as hybridization probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA.

Claim 1; SEQ ID NO 35150; 44pp; English.

The invention relates to an isolated polynucleotide comprising any one of 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was determined by the technique of SBH (sequencing by hybridisation). Also included is a purified polypeptide comprising a sequence corresponding to a reading frame of the novel polynucleotide. The nucleic acid sequences are useful in diagnostics as expressed sequence tags (EST) for identifying expressed genes or for physical mapping of the human genome, in forensic, in assessing biodiversity, or in identifying mutations responsible for genetic disorders and other traits. The nucleotide sequences are also useful as hybridisation probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA. The purified polypeptide is useful for generating antibodies specific for it. The present sequence data is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030073623



Qy 1 ATGCACCTTCTGTGGAGGCACCTTGATATATCCCGAGAGTGGGTGTGACTGCTGCCACTGC 60  
Db 255 ATGCACCTTCTGTGGAGGCACCTTGATATATCCCGAGAGTGGGTGTGACTGCTGCCACTGC 314  
Qy 61 TTGGAGAGTCCCGAAGGCCCTTCATCCTACAAAGGTATCTCTGGGTGC 107  
Db 315 TTGGAGAGTCCCGAAGGCCCTTCATCCTACAAAGGTATCTCTGGGTGC 361

RESULT 8	
ACH42960	
ID	ACH42960 standard; cDNA; 357 BP.
XX	
XX	
ACH42960;	
XX	
XX	
DT	13-OCT-2003 (first entry)
XX	
XX	
DE	Human foetal liver/spleen cDNA #176.
XX	
XX	
KW	Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW	genome mapping; biodiversity; genetic disorder.

```
Query Match      16.6%; Score 103; DB 9; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.5e-43;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

1 ATGCACCTTCGTGGAGGACCTTGATATCCCCAGAGTGGGTGTTGACTGCTGCCACTGC 60  
255 ATGCACCTTCGTGGAGGACCTTGATATCCCCAGAGTGGGTGTTGACTGCTGCCACTGC 314

61 TTGGAGAGTCCCAAGGCGCTTCATCCTACAGGTGATCTCTGG 103  
315 TTGGAGAGTCCCAAGGCGCTTCATCCTACAGGTGATCTCTGG 357

Qy Db Qy Db Qy Db

RESULT 9	
AAL26661	
ID	AAL26661 standard; cDNA; 468 BP.
XX	
XX	
AC	AAL26661;
XX	
XX	07-DEC-2001 (first entry)
DT	
XX	
XX	Human breast cancer expressed polynucleotide 19118.
DE	
XX	
XX	Human; breast cancer; cell marker; cytostatic; ss.
KW	
XX	
XX	Homo sapiens.
XX	
XX	WO200151628-A2.
PN	
XX	
PD	19-JUL-2001.
XX	
XX	
PF	10-JAN-2001; 2001WO-US000798.
XX	
XX	14-JAN-2000; 2000US-0176077P.
PR	
PR	14-MAR-2000; 2000US-0189167P.
PR	
PR	24-MAR-2000; 2000US-0192099P.
PR	
PR	29-MAR-2000; 2000US-0193480P.
PR	
PR	15-MAY-2000; 2000US-0205230P.
PR	
PR	09-JUN-2000; 2000US-0211315P.
PR	
PR	25-JUL-2000; 2000US-0220534P.
XX	
XX	
PA	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX	
PI	Lillie J, Xu Y, Wang Y, Steinmann K;
XX	
XX	WPI; 2001-451856/48.
DR	

	Query Match	5.6%	Score 35;	DB 4;	Length 468;
Best Local Similarity	100.0%				
Matches 35; Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;				
QY	357	TGTGATTGAGAATAAAGTGTGCAATCGCTATGAGT	391		
DB	40	TGTGATTGAGAATAAAGTGTGCAATCGCTATGAGT	74		

RESULT 10  
ABX47578  
ID ABX47578 standard; CDNA; 346 BP.  
XX  
AC ABX47578;

XX	DT	21-FEB-2003 (first entry)	XX	DT	21-FEB-2003 (first entry)			
XX	DE	Bovine EST associated with lactation/muscle/fat deposition #12743.	XX	DE	ACH20439 standard; cDNA; 384 BP.			
XX	KW	Bovine; ss; EST; expressed sequence tag; lactation; LMFD; muscle deposition; fat deposition; genome mapping; gene identification; gene analysis; cattle breeding.	XX	KW	ACH20439; AC			
XX	OS	Bos Taurus.	XX	OS	ACH20439; AC			
XX	PN	US2002137139-A1.	XX	PN	13-OCT-2003 (first entry)			
XX	PD	26-SEP-2002.	XX	PD	Human adult liver cDNA #51.			
XX	PP	24-SEP-2001; 2001US-00960352.	XX	PP	Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST; genome mapping; biodiversity; genetic disorder.			
XX	PR	12-JAN-1999; 99US-0115707P.	XX	PR	Homo sapiens.			
XX	PR	11-JAN-2000; 2000US-00480902.	XX	PR	US2003073623-A1.			
XX	PA	(BYAT/) BYATT J C.	XX	PA	17-APR-2003.			
XX	PA	(MATH/) MATHIALAGAN N.	XX	PA	30-JUL-2001; 2001US-00918995.			
XX	PA	(TAON/) TAO N.	XX	PA	30-JUL-2001; 2001US-00918995.			
XX	PA	(WARR/) WARREN W C.	XX	PA	(DRMA/) DRMANAC R T.			
XX	PI	Byatt JC, Mathialagan N, Tao N, Warren WC;	XX	PI	(LABA/) LABAT I.			
XX	DR	WPI; 2003-110599/10.	XX	DR	(STAC/) STACHE-CRAIN B.			
XX	XX	New nucleic acid associated with lactation, and muscle and fat deposition, useful for genome mapping, gene identification and analysis, cattle breeding, or for genetically improving cattle.	XX	XX	(DICK/) DICKSON M C.			
XX	PS	Claim 2; SEQ ID NO 12743; 245pp; English.	XX	PS	(JONE/) JONES L W.			
XX	CC	The invention relates to a purified nucleic acid molecule associated with lactation or muscle and fat deposition (designated LMFD), derived from cattle, and the LMFD nucleic acid can specifically hybridise to a second nucleic acid molecule comprising any of 15112 nucleotide sequences, appearing as ABX34836-ABX49947, or complements of them. Also included are (1) a transformed cell having a nucleic acid comprising an LMFD nucleic acid linked to a promoter and a 3' non-translated sequence that functions in the cell to cause termination of transcription and addition of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and (2) determining a level or pattern of a molecule in a bovine cell or tissue comprising: (a) incubating a marker nucleic acid (comprising any of the 15112 nucleic acid sequences or its complement or fragment) with a complementary nucleic acid molecule obtained from the bovine cell or tissue, where hybridisation between the marker nucleic acid and the complementary nucleic acid permits the detection of the molecule; and (b) detecting the level or pattern of the complementary nucleic acid, where the detection of the complementary nucleic acid is predictive of the level or pattern of the molecule. The LMFD nucleic acid is used for determining a level or pattern of a molecule in a bovine cell or tissue. It is useful for genome mapping, gene identification and analysis, cattle breeding, preparation of constructs for use in cattle gene expression, or for genetically improving cattle. The present sequence is one of the 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present sequence was not shown in the specification but was obtained in electronic format from the USPTO web site: seqdata.uspto.gov/sequence.html?DocID=20020137139	XX	CC	Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW; WPI; 2003-615964/58.			
XX	CC	Sequence 346 BP; 79 A; 88 C; 107 G; 72 T; 0 U; 0 Other;	XX	CC	New polynucleotide sequences obtained from various cDNA libraries, useful as hybridization probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA.			
XX	CC	Query Match 5.3%; Score 33; DB 8; Length 346;	XX	CC	Claim 1; SEQ ID NO 7651; 44pp; English.			
XX	CC	Best Local Similarity 100.0%; Pred. No. 1.1e-06;	XX	CC	The invention relates to an isolated polynucleotide comprising any one of 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was determined by the technique of SBH (sequencing by hybridisation). Also included is a purified polypeptide comprising a sequence corresponding to a reading frame of the novel polynucleotide. The nucleic acid sequences are useful in diagnostics as expressed sequence tags (EST) for identifying expressed genes or for physical mapping of the human genome, in forensics, in assessing biodiversity, or in identifying mutations responsible for genetic disorders and other traits. The nucleotide sequences are also useful as hybridisation probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA. The purified polypeptide is useful for generating antibodies specific for it. The present sequence is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030073623	XX	CC	Sequence 384 BP; 94 A; 95 C; 110 G; 84 T; 0 U; 1 Other;
XX	CC	Sequence 346 BP; 79 A; 88 C; 107 G; 72 T; 0 U; 0 Other;	XX	CC	Query Match 4.8%; Score 30; DB 9; Length 384;			
XX	CC	Best Local Similarity 100.0%; Pred. No. 1.1e-06;	XX	CC	Best Local Similarity 100.0%; Pred. No. 4.4e-05;			
XX	CC	Mismatches 0; Indels 0; Gaps 0;	XX	CC	Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	CC	Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Sequence 346 BP; 79 A; 88 C; 107 G; 72 T; 0 U; 0 Other;	XX	CC	Sequence 384 BP; 94 A; 95 C; 110 G; 84 T; 0 U; 1 Other;			
XX	CC	Query Match 5.3%; Score 33; DB 8; Length 346;	XX	CC	Query Match 4.8%; Score 30; DB 9; Length 384;			
XX	CC	Best Local Similarity 100.0%; Pred. No. 1.1e-06;	XX	CC	Best Local Similarity 100.0%; Pred. No. 4.4e-05;			
XX	CC	Mismatches 0; Indels 0; Gaps 0;	XX	CC	Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	CC	Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Sequence 346 BP; 79 A; 88 C; 107 G; 72 T; 0 U; 0 Other;	XX	CC	Sequence 384 BP; 94 A; 95 C; 110 G; 84 T; 0 U; 1 Other;			
XX	CC	Query Match 5.3%; Score 33; DB 8; Length 346;	XX	CC	Query Match 4.8%; Score 30; DB 9; Length 384;			
XX	CC	Best Local Similarity 100.0%; Pred. No. 1.1e-06;	XX	CC	Best Local Similarity 100.0%; Pred. No. 4.4e-05;			
XX	CC	Mismatches 0; Indels 0; Gaps 0;	XX	CC	Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	CC	Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Sequence 346 BP; 79 A; 88 C; 107 G; 72 T; 0 U; 0 Other;	XX	CC	Sequence 384 BP; 94 A; 95 C; 110 G; 84 T; 0 U; 1 Other;			
XX	CC	Query Match 5.3%; Score 33; DB 8; Length 346;	XX	CC	Query Match 4.8%; Score 30; DB 9; Length 384;			
XX	CC	Best Local Similarity 100.0%; Pred. No. 1.1e-06;	XX	CC	Best Local Similarity 100.0%; Pred. No. 4.4e-05;			
XX	CC	Mismatches 0; Indels 0; Gaps 0;	XX	CC	Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	XX	CC	Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
XX	CC	Sequence 346 BP; 79 A; 88 C; 107 G; 72 T; 0 U; 0 Other;	XX	CC	Sequence 384 BP; 94 A; 95 C; 110 G; 84 T; 0 U; 1 Other;			
XX	CC	Query Match 5.3%; Score 33; DB 8; Length 346;	XX	CC	Query Match 4.8%; Score 30; DB 9; Length 384;			
XX	CC	Best Local Similarity 10						



```

PR 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
PI WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 11088; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid molecule permits the detection of the molecule; and
CC (b) detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
XX Sequence 394 BP; 100 A; 104 C; 119 G; 71 T; 0 U; 0 Other;
SQ
Query Match 4.3%; Score 27; DB 8; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 CATCACTGGCTGGGAGAAACCCCAAGG 317
DB 202 CATCACTGGCTGGGAGAAACCCCAAGG 228

RESULT 15
ABX45686
ID ABX45686 standard; cDNA; 468 BP.
XX
XX ABX45686;
AC
XX
XX 21-FEB-2003 (first entry)
DT
XX
XX Bovine EST associated with lactation/muscle/fat deposition #10851.
DE
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
XX Bos Taurus.
OS
XX
XX US2002137139-A1.
PN
XX

```

---

```

PD 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
PI WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 10851; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and
CC (b) detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
XX Sequence 468 BP; 111 A; 129 C; 137 G; 91 T; 0 U; 0 Other;
SQ
Query Match 4.3%; Score 27; DB 8; Length 468;
Best Local Similarity 100.0%; Pred. No. 0.0017;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 CATCACTGGCTGGGAGAAACCCCAAGG 317
DB 245 CATCACTGGCTGGGAGAAACCCCAAGG 271

Search completed: October 27, 2004, 06:00:17
Job time : 463 secs

```



GenCore version 5.1.16  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 04:11:05 ; Search time 101 Seconds  
(without alignments)  
4370.236 Million cell updates/sec

Title: US-09-992-095b-53\_COPY\_1044\_1664  
Perfect score: 621  
Sequence: 1 atgcactctgtggaggcac.....aggagtgatgagaataat 621

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 824507 seqs, 355394441 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1390726

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	13.4	312	4	US-09-621-976-9524
2	20	3.2	39	1	US-08-456-840-11
3	20	3.2	39	1	US-08-456-840-15
4	20	3.2	39	1	US-08-266-407A-11
5	20	3.2	39	1	US-08-266-407A-15
6	20	3.2	39	2	US-08-892-544-11
7	20	3.2	39	2	US-08-892-544-15
8	19	3.1	36	1	US-08-456-840-19
9	19	3.1	36	1	US-08-266-407A-19
10	19	3.1	36	2	US-08-892-544-19
11	19	3.1	39	1	US-08-456-840-3
12	19	3.1	39	1	US-08-456-840-9
13	19	3.1	39	1	US-08-266-407A-3
14	19	3.1	39	1	US-08-266-407A-9
15	19	3.1	39	2	US-08-892-544-3
16	19	3.1	39	2	US-08-892-544-9
17	18	2.9	39	1	US-08-456-840-13
18	18	2.9	39	1	US-08-266-407A-13
19	18	2.9	39	2	US-08-892-544-13
20	18	2.9	361	4	US-09-513-999C-13535
21	17	2.7	411	4	US-09-854-133-361
22	17	2.7	447	4	US-09-513-999C-36190
23	17	2.7	481	3	US-09-280-116-155
24	16	2.6	16	1	US-08-311-760A-352
25	16	2.6	16	1	US-08-311-760A-353
26	16	2.6	16	2	US-08-774-310-352
27	16	2.6	16	2	US-08-774-310-353

28	16	2.6	33	1	US-08-456-840-1	Sequence 1, Appli
29	16	2.6	33	1	US-08-266-407A-1	Sequence 1, Appli
30	16	2.6	33	2	US-08-892-544-1	Sequence 1, Appli
31	16	2.6	36	1	US-08-456-840-21	Sequence 21, Appl
32	16	2.6	36	1	US-08-266-407A-21	Sequence 21, Appl
33	16	2.6	36	2	US-08-892-544-21	Sequence 21, Appl
34	16	2.6	201	4	US-09-248-796A-3253	Sequence 3253, Ap
35	16	2.6	212	4	US-09-513-999C-32078	Sequence 32078, A
36	16	2.6	318	4	US-09-270-767-10991	Sequence 10991, A
37	16	2.6	333	4	US-09-248-796A-7950	Sequence 7950, A
38	16	2.6	406	4	US-09-621-976-19198	Sequence 19198, A
39	16	2.6	447	4	US-09-621-976-10672	Sequence 10672, A
40	16	2.6	453	4	US-09-543-681A-3641	Sequence 3641, Ap
41	16	2.6	461	4	US-09-270-767-5257	Sequence 5257, Ap
42	16	2.6	461	4	US-09-270-767-20539	Sequence 20539, A
43	16	2.6	498	4	US-09-252-991A-2919	Sequence 2919, Ap
44	16	2.6	609	4	US-09-543-681A-3680	Sequence 3680, Ap
45	15	2.4	19	4	US-09-696-791-3201	Sequence 3201, Ap

ALIGNMENTS

RESULT 1  
US-09-621-976-9524  
; Sequence 9524, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 9524  
; LENGTH: 312  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 189  
; OTHER INFORMATION: n=a, g, c or t  
; US-09-621-976-9524

Query Match	13.4%	Score 83;	DB 4;	Length 312;
Best Local Similarity	100.0%;	Pred. No. 7.7e-33;		
Matches 83;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	363	TGAGATAAGTGTGCAATCGGTATGAGTTTCTGAATGAAGAGTCCCAATCCACCGAACT	422	
Db	1	TGAGATAAGTGTGCAATCGGTATGAGTTTCTGAATGAAGAGTCCCAATCCACCGAACT	60	
QY	423	CTGTGCTGGGCATTTGGCCGGAG	445	
Db	61	CTGTGCTGGGCATTTGGCCGGAG	83	

RESULT 2  
US-08-456-840-11  
; Sequence 11, Application US/08456840  
; Patent No. 5597908  
; GENERAL INFORMATION:  
; APPLICANT: Taddei-Peters, W. C.  
; APPLICANT: Butler, Sandra M.  
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5597908el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville

```

; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,840
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
; US-08-456-840-11

Query Match 3.2%; Score 20; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 CAAGGCCTTCATCCTACAG 93
Db 20 CAAGGCCTTCATCCTACAG 39

RESULT 3
US-08-456-840-15
; Sequence 15, Application US/08456840
; Patent No. 5597908
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5597908el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,840
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/266,407
; FILING DATE: 27-JUN-1994

; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,840
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/266,407
; FILING DATE: 27-JUN-1994

```

```

; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
; US-08-456-840-15

Query Match 3.2%; Score 20; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 GCCTTGCTAAAGCTAAGCAG 209
Db 1 GCCTTGCTAAAGCTAAGCAG 20

RESULT 4
US-08-266-407A-11
; Sequence 11, Application US/08266407A
; Patent No. 5786156
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5786156el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407A
; FILING DATE: 27-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: 1..39
US-08-266-407A-11

Query Match          3.2%; Score 20; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 CAAGGCTTCATCCTACAAG 93
Db 20 CAAGGCTTCATCCTACAAG 39

RESULT 5
US-08-266-407A-15
; Sequence 15, Application US/08266407A
; Patent No. 5786156
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5786156el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
US-08-892-544-11

Query Match          3.2%; Score 20; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 CAAGGCTTCATCCTACAAG 93
Db 20 CAAGGCTTCATCCTACAAG 39

RESULT 7
US-08-892-544-15
; Sequence 15, Application US/08892544
; Patent No. 5874544
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5874544el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407A
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
US-08-266-407A-15

Query Match          3.2%; Score 20; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 190 GCCTTGCTAAAGCTAACGAC 209
Db 1 GCCTTGCTAAAGCTAACGAC 20

RESULT 6
US-08-892-544-11
; Sequence 11, Application US/08892544
; Patent No. 5874544
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/892,544  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/266,407  
 FILING DATE: 27-JUN-1994  
 APPLICATION NUMBER: US/08/172,461  
 FILING DATE: 21-DEC-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gormley, Mary E.  
 REGISTRATION NUMBER: 34,409  
 TELEPHONE: 301-258-5200  
 TELEFAX: 301-977-0847  
 INFORMATION FOR SEQ ID NO: 15:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 39 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..39  
 US-08-892-544-15

Query Match 3.2%; Score 20; DB 2; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 190 GCCTTGCTAAAGCTAAGCAG 209  
 Db 1 GCCTTGCTAAAGCTAAGCAG 20

RESULT 8  
 US-08-456-840-19  
 Sequence 19, Application US/08456840  
 Patent No. 5597908  
 GENERAL INFORMATION:  
 APPLICANT: Taddei-Peters, W. C.  
 APPLICANT: Butler, Sandra M.  
 TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
 NUMBER OF SEQUENCES: 48  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Akzo No. 5597908el  
 STREET: 1330 Piccard Drive  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: US  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/456,840  
 FILING DATE: 01-JUN-1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/266,407  
 FILING DATE: 27-JUN-1994  
 APPLICATION NUMBER: US/08/172,461  
 FILING DATE: 21-DEC-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gormley, Mary E.  
 REGISTRATION NUMBER: 34,409  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-258-5200  
 TELEFAX: 301-977-0847  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 36 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..36  
 US-08-456-840-19

Query Match 3.1%; Score 19; DB 1; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 CTGTTCTTGGAGCCACAC 178  
 Db 1 CTGTTCTTGGAGCCACAC 19

RESULT 9  
 US-08-266-407A-19  
 Sequence 19, Application US/08266407A  
 Patent No. 5786156  
 GENERAL INFORMATION:  
 APPLICANT: Taddei-Peters, W. C.  
 APPLICANT: Butler, Sandra M.  
 TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
 NUMBER OF SEQUENCES: 48  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Akzo No. 5786156el  
 STREET: 1330 Piccard Drive  
 CITY: Rockville  
 STATE: Maryland  
 COUNTRY: US  
 ZIP: 20850  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/266,407A  
 FILING DATE: 27-JUN-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/172,461  
 FILING DATE: 21-DEC-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gormley, Mary E.  
 REGISTRATION NUMBER: 34,409  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 301-258-5200  
 TELEFAX: 301-977-0847  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 36 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: both  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..36  
 US-08-266-407A-19

Query Match 3.1%; Score 19; DB 1; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 4.9;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 CTGTTCTTGGAGCCACAC 178  
Db 1 CTGTTCTTGGAGCCACAC 19

RESULT 10  
US-08-892-544-19  
; Sequence 19, Application US/08892544  
; Patent No. 5874544  
; GENERAL INFORMATION:  
; APPLICANT: Taddei-Peters, W. C.  
; APPLICANT: Butler, Sandra M.  
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5874544el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/892,544  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/266,407  
; FILING DATE: 27-JUN-1994  
; APPLICATION NUMBER: US 08/172,461  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gormley, Mary E.  
; REGISTRATION NUMBER: 34,409  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-258-5200  
; TELEFAX: 301-977-0847  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 36 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..36  
US-08-892-544-19  
Query Match 3.1%; Score 19; DB 2; Length 36;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 160 CTGTTCTTGGAGCCACAC 178  
Db 1 CTGTTCTTGGAGCCACAC 19  
RESULT 11  
US-08-456-840-3  
; Sequence 3, Application US/08456840  
; Patent No. 5597908  
; GENERAL INFORMATION:  
; APPLICANT: Taddei-Peters, W. C.  
; APPLICANT: Butler, Sandra M.  
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5597908el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,840

; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5597908el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,840  
; FILING DATE: 01-JUN-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/266,407  
; FILING DATE: 27-JUN-1994  
; APPLICATION NUMBER: US 08/172,461  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gormley, Mary E.  
; REGISTRATION NUMBER: 34,409  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 301-258-5200  
; TELEFAX: 301-977-0847  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 39 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..39  
US-08-456-840-3  
Query Match 3.1%; Score 19; DB 1; Length 39;  
Best Local Similarity 100.0%; Pred. No. 4.9;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 291 CATCACTGGCTGGGAGAA 309  
Db 21 CATCACTGGCTGGGAGAA 39  
RESULT 12  
US-08-456-840-9  
; Sequence 9, Application US/08456840  
; Patent No. 5597908  
; GENERAL INFORMATION:  
; APPLICANT: Taddei-Peters, W. C.  
; APPLICANT: Butler, Sandra M.  
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5597908el  
; STREET: 1330 Piccard Drive  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: US  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,840

```

; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
US-08-456-840-9

```

Query Match 3.1%; Score 19; DB 1; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 GCACCTCTGTGGAGCACC 21
    |||||
DB 21 GCACCTCTGTGGAGCACC 39

```

```

RESULT 13
US-08-266-407A-3
; Sequence 3, Application US/08266407A
; Patent No. 5786156
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5786156el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407A
; FILING DATE: 27-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid

```

```

; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
US-08-266-407A-3

```

Query Match 3.1%; Score 19; DB 1; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 291 CATCACTGGCTGGGAGAA 309
    |||||
DB 21 CATCACTGGCTGGGAGAA 39

```

```

RESULT 14
US-08-266-407A-9
; Sequence 9, Application US/08266407A
; Patent No. 5786156
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5786156el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407A
; FILING DATE: 27-JUN-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
US-08-266-407A-9

```

Query Match 3.1%; Score 19; DB 1; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 GCACCTCTGTGGAGCACC 21
    |||||
DB 21 GCACCTCTGTGGAGCACC 39

```

```

RESULT 15
US-08-892-544-3
; Sequence 3, Application US/08892544
; Patent No. 5874544
; GENERAL INFORMATION:
; APPLICANT: Taddei-Peters, W. C.
; APPLICANT: Butler, Sandra M.
; TITLE OF INVENTION: Immunoreactive Peptides of Apo(a)
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5874544el
; STREET: 1330 Piccard Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892,544
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/266,407
; FILING DATE: 27-JUN-1994
; APPLICATION NUMBER: US 08/172,461
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gormley, Mary E.
; REGISTRATION NUMBER: 34,409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-258-5200
; TELEFAX: 301-977-0847
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..39
; US-08-892-544-3

```

```

Query Match 3.1%; Score 19; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 291 CATCACTGGCTGGGGAGAA 309
Db 21 CATCACTGGCTGGGGAGAA 39

```

```

Search completed: October 27, 2004, 07:39:14
Job time : 103 secs

```

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 06:51:10 ; Search time 1644 Seconds  
(without alignments)  
1936.886 Million cell updates/sec

Title: US-09-992-095b-53\_COPY\_1044\_1664

Perfect score: 621

Sequence: 1 atgcactctctggaggcac.....aggagtgatgagaataat 621

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3413475 seqs, 2563800928 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4670548

Minimum DB seq length: 0

Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	24.2	201	17	US-10-741-601-2633
2	150	24.2	201	17	US-10-741-601-2639
3	150	24.2	201	17	US-10-741-601-2649
4	150	24.2	201	17	US-10-741-601-2657
5	150	24.2	201	17	US-10-741-601-2657
6	129	20.8	508	10	US-09-918-995-35150
7	107	17.2	363	10	US-09-918-995-30128
8	103	16.6	357	10	US-09-918-995-30172
9	100	16.1	201	17	US-10-741-601-2629
10	100	16.1	201	17	US-10-741-601-2660
11	100	16.1	201	17	US-10-741-601-14376
12	100	16.1	201	17	US-10-741-601-14415
					Sequence 2633, Ap
					Sequence 2639, Ap
					Sequence 2649, Ap
					Sequence 2653, Ap
					Sequence 2657, Ap
					Sequence 35150, A
					Sequence 30128, A
					Sequence 30172, A
					Sequence 2629, Ap
					Sequence 2660, Ap
					Sequence 14376, A
					Sequence 14415, A

13	93	15.0	201	17	US-10-741-601-2630	Sequence 2630, Ap
14	93	15.0	201	17	US-10-741-601-14280	Sequence 14280, A
15	90	14.5	201	17	US-10-741-601-14344	Sequence 14344, A
16	84	13.5	201	17	US-10-741-601-14453	Sequence 14453, A
17	74	11.9	201	17	US-10-741-601-14495	Sequence 14495, A
18	67	10.8	201	17	US-10-741-601-14401	Sequence 14401, A
19	66	10.6	201	17	US-10-741-601-14432	Sequence 14432, A
20	64	10.3	201	17	US-10-741-601-14486	Sequence 14486, A
21	53	8.5	201	17	US-10-741-601-2659	Sequence 2659, Ap
22	53	8.5	201	17	US-10-741-601-14374	Sequence 14374, A
23	51	8.2	201	17	US-10-741-601-14327	Sequence 14327, A
24	48	7.7	201	17	US-10-741-601-14445	Sequence 14445, A
25	45	7.2	201	17	US-10-741-601-14420	Sequence 14420, A
26	33	5.3	346	9	US-09-960-352-12743	Sequence 12743, A
27	30	4.8	384	10	US-09-918-995-7651	Sequence 7651, Ap
28	27	4.3	394	9	US-09-960-352-11088	Sequence 11088, A
29	27	4.3	468	9	US-09-960-352-10851	Sequence 10851, A
30	26	4.2	299	11	US-09-864-408A-5459	Sequence 5459, Ap
31	25	4.0	162	9	US-09-960-352-7082	Sequence 7082, Ap
32	25	4.0	203	9	US-09-960-352-11298	Sequence 11298, A
33	25	4.0	234	9	US-09-960-352-14383	Sequence 14383, A
34	21	3.4	38	16	US-10-450-976-8	Sequence 8, Appli
35	21	3.4	384	9	US-09-960-352-10993	Sequence 10993, A
36	20	3.2	41	16	US-10-453-827-162	Sequence 162, App
37	18	2.9	165	15	US-10-029-386-15124	Sequence 15124, A
38	18	2.9	221	15	US-10-029-386-21008	Sequence 21008, A
39	18	2.9	324	9	US-09-983-965-3394	Sequence 3394, Ap
40	18	2.9	344	16	US-10-424-599-78352	Sequence 78352, A
41	18	2.9	421	10	US-09-918-995-5157	Sequence 5157, Ap
42	18	2.9	491	10	US-09-918-995-19684	Sequence 19684, A
43	18	2.9	523	15	US-10-029-386-1202	Sequence 1202, Ap
44	18	2.9	540	15	US-10-029-386-1422	Sequence 1422, Ap
45	18	2.9	569	15	US-10-029-386-7308	Sequence 7308, Ap

ALIGNMENTS

RESULT 1

US-10-741-601-2633  
; Sequence 2633, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2633  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-2633

Query Match 24.2%; Score 150; DB 17; Length 201;

Best Local Similarity 99.5%; Pred. No. 28-73; 1; Indels 0; Gaps 0;  
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	32	CAGAGTGGGTGTGACTGCTGCCACTGCTTGGAGAAGTCCCAAGGCCCTTCACTCTACA	91
Db	1	CAGAGTGGGTGTGACTGCTGCCACTGCTTGGAGAAGTCCCAAGGCCCTTCACTCTACA	60
QY	92	AGGTATCTCGGTGCACACCAAGAGTGAATCTCGAAGCGCATGTTTCAGGAATAGAAG	151
Db	61	AGGTATCTCGGTGCACACCAAGAGTGAATCTCGAAGCGCATGTTTCAGGAATAGAAG	120
QY	152	TGCTTAGGCTGTTCTTGGAGCCACACGAAAGATATTGCTTGTAAAGCTAAGCAGTC	211
Db	121	TGCTTAGGCTGTTCTTGGAGCCACACGAAAGATATTGCTTGTAAAGCTAAGCAGTC	180

```

QY 212 CTGCGTCACTGACAAAG 232
DB 181 CTGCGTCACTGACAAAG 201

RESULT 2
US-10-741-601-2639
; Sequence 2639, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2639
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-2639

Query Match 24.2%; Score 150; DB 17; Length 201;
Best Local Similarity 99.5%; Pred. No. 2e-73;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 174 CACAGAAAGATATTGCTTAAAGTAAGCAGTCTGCGCTCATCACTGACAAAGT 233
DB 1 CACAGAAAGATATTGCTTAAAGTAAGCAGTCTGCGCTCATCACTGACAAAGT 60

QY 234 AATCCAGCTTGTGCCATCCCCAAATTAATGTGCTGACCGGACCGAATGTTTCAT 293
DB 61 AATCCAGCTTGTGCCATCCCCAAATTAATGTGCTGACCGGACCGAATGTTTCAT 120

QY 294 CACTGGCTGGGAGAAACCAAGTACTTTGGAGCTGCGCTTCAAGGAAGCCAGCT 353
DB 121 CACTGGCTGGGAGAAACCAAGTACTTTGGAGCTGCGCTTCAAGGAAGCCAGCT 180

QY 354 CCCTGTGATTGAGATAAAGT 374
DB 181 CCCTGTGATTGAGATAAAGT 201

RESULT 3
US-10-741-601-2649
; Sequence 2649, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2649
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-2649

Query Match 24.2%; Score 150; DB 17; Length 201;
Best Local Similarity 99.5%; Pred. No. 2e-73;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 377 GCAATCGCTATGAGTTTCTGAATGAAGAGTCCAAATCCACCGAATCTGTGCTGGGCATT 436
DB 1 GCAATCGCTATGAGTTTCTGAATGAAGAGTCCAAATCCACCGAATCTGTGCTGGGCATT 60

QY 437 TGGCCGGAGGCACTGACAGTTGCCAGGTGACAGTGGAGGTCCTCTGTTGCTTCGAGA 496

```

---

```

DB 61 TGGCCGGAGGCACTGACAGTTGCCAGGTGACAGTGGAGGKCCCTCTGTTGCTTCGAGA 120
QY 497 AGGACAAATACATTTTACAAGGAGTCACTTCTTGGGCTCTTGGCTGTGACAGCCCAATA 556
DB 121 AGGACAAATACATTTTACAAGGAGTCACTTCTTGGGCTCTTGGCTGTGACAGCCCAATA 180

QY 557 AGCCTGTGTCTATGTTTCGTG 577
DB 181 AGCCTGTGTCTATGTTTCGTG 201

RESULT 4
US-10-741-601-2653
; Sequence 2653, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2653
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-2653

Query Match 24.2%; Score 150; DB 17; Length 201;
Best Local Similarity 99.5%; Pred. No. 2e-73;
Matches 200; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 171 GCCCACAGAAAGATATTGCTTAAAGTAAGCAGTCTGCGCTCATCACTGACAA 230
DB 1 GCCCACAGAAAGATATTGCTTAAAGTAAGCAGTCTGCGCTCATCACTGACAA 60

QY 231 AGTAATCCAGCTTGTCTGCCATCCCCAAATTAATGTGCTGACCGGACCGAATGTTT 290
DB 61 AGTAATCCAGCTTGTCTGCCATCCCCAAATTAATGTGCTGACCGGACCGAATGTTT 120

QY 291 CATCACTGGCTGGGAGAAACCAAGTACTTTTGGAGCTGCGCTTCAAGGAAGCCCA 350
DB 121 CATCACTGGCTGGGAGAAACCAAGTACTTTTGGAGCTGCGCTTCAAGGAAGCCCA 180

QY 351 GCTCCCTGTGATTGAGAATAA 371
DB 181 GCTCCCTGTGATTGAGAATAA 201

RESULT 5
US-10-741-601-2657
; Sequence 2657, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2657
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-2657

Query Match 24.2%; Score 150; DB 17; Length 201;
Best Local Similarity 99.5%; Pred. No. 2e-73;

```

	Query Match	17.2%	Score 107;	DB 10;	Length 363;
	Best Local Similarity	100.0%;	Pred. No. 3.2e-49;		
	Matches 107;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGCAC	TTCTGTGGAGGCAC	TTTGATATCCCCAGAGTGGGTGTGTGACTGTGCCCCATGC	60
Db	255	ATGCAC	TTCTGTGGAGGCAC	TTTGATATCCCCAGAGTGGGTGTGTGACTGTGCCCCATGC	314
Qy	61	TTGGAG	AAAGTCCCAAGGCC	TTTCATCTCAAGGTATCTCTGGGTGC	107
Db	315	TTGGAG	AAAGTCCCAAGGCC	TTTCATCTCAAGGTATCTCTGGGTGC	361

```

RESULT 8
US-09-918-995-30172
; Sequence 30172, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30172
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-918-995-30172

```

	Query Match	16.6%	Score 103;	DB 10;	Length 357;
	Best Local Similarity	100.0%;	Pred. No. 5.7e-47;		
	Matches 103; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	ATGCAC TTC TGTGAGGAC CTTTCAT ATATCC CAGAGTGG GTGTGACTG CTGCCACATGC	60		
Dd	255	ATGCAC TTC TGTGAGGAC CTTTCAT ATATCC CAGAGTGG GTGTGACTG CTGCCACATGC	314		
Qy	61	TTGAGA AGTCCCCA AGCCTTC ATCTCA AAGTCAT CCTCGG	103		
Dd	315	TTGAGA AGTCCCCA AGCCTTC ATCTCA AAGTCAT CCTCGG	357		

```

RESULT 9
US-10-741-601-2629
; Sequence 2629, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: Fast-SEQ for Windows Version 4.0

```

```
; SEQ ID NO 2629
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-2629

Query Match      16.1%; Score 100; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-45;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 CTGCCACTGTGTGAGAGTCCCAAGGCTTCATCTACAAGTCAATCTCGGTGTCAC 109
Db 102 CTGCCACTGTGTGAGAGTCCCAAGGCTTCATCTACAAGTCAATCTCGGTGTCAC 161

Qy 110 ACCAAGAGTGAATCTCGAACCGCATGTTTCAGGAATAAGA 149
Db 162 ACCAAGAGTGAATCTCGAACCGCATGTTTCAGGAATAAGA 201

RESULT 10
US-10-741-601-2660
; Sequence 2660, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2660
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-2660

Query Match      16.1%; Score 100; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-45;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 350 AGCTCCCTGTGATGAGATAAAGTGCATCGTATGAGTTTCTGAATGGAAGTCC 409
Db 1 AGCTCCCTGTGATGAGATAAAGTGCATCGTATGAGTTTCTGAATGGAAGTCC 60

Qy 410 AATCCACCGAACTCTGTGCTGGCATTGGCCGGAGGCAC 449
Db 61 AATCCACCGAACTCTGTGCTGGCATTGGCCGGAGGCAC 100

RESULT 11
US-10-741-601-14376
; Sequence 14376, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14376
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-14376

Query Match      16.1%; Score 100; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-45;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 350 AGCTCCCTGTGATGAGATAAAGTGCATCGTATGAGTTTCTGAATGGAAGTCC 409
Db 1 AGCTCCCTGTGATGAGATAAAGTGCATCGTATGAGTTTCTGAATGGAAGTCC 60

Qy 410 AATCCACCGAACTCTGTGCTGGCATTGGCCGGAGGCAC 449
Db 61 AATCCACCGAACTCTGTGCTGGCATTGGCCGGAGGCAC 100

RESULT 12
US-10-741-601-14415
; Sequence 14415, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14415
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-14415

Query Match      16.1%; Score 100; DB 17; Length 201;
Best Local Similarity 100.0%; Pred. No. 2.7e-45;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 478 CCTCTGGTTTGTCTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTT 537
Db 102 CCTCTGGTTTGTCTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTT 161

Qy 538 GGCTGTGCACGCCCAATAGCCCTGCTCTATGTTCTG 577
Db 162 GGCTGTGCACGCCCAATAGCCCTGCTCTATGTTCTG 201

RESULT 13
US-10-741-601-2630
; Sequence 2630, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2630
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-2630

Query Match      15.0%; Score 93; DB 17; Length 201;
Best Local Similarity 99.3%; Pred. No. 2.3e-41;
Matches 143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 478 CCTCTGGTTTGTCTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTT 537
Db 18 CCTCTGGTTTGTCTCGAGAAGGACAAATACATTTTACAAGGAGTCACTTCTTGGGGTCTT 77

Qy 538 GGCTGTGCACGCCCAATAGCCCTGCTCTATGTTCTGTTTCAAGTTTGTACTTGG 597
Db 78 GGCTGTGCACGCCCAATAGCCCTGCTCTATGTTCTGTTTCAAGTTTGTACTTGG 137
```

QY 598 ATTGAGGAGTGATGAGAAATAAT 621  
|||||  
Db 138 ATTGAGGAGTGATGAGAAATAAT 161  
|||||  
Search completed: October 27, 2004, 09:04:42  
Job time : 1646 secs

QY 598 ATTGAGGAGTGATGAGAAATAAT 621  
|||||  
Db 138 ATTGAGGAGTGATGAGAAATAAT 161  
|||||

RESULT 14  
US-10-741-601-14280  
; Sequence 14280, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 14280  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-14280

Query Match 15.0%; Score 93; DB 17; Length 201;  
Best Local Similarity 99.3%; Pred. No. 2.3e-41;  
Matches 143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 478 CCTCTGGTTTGTTCGAGAGGACAAATACATTTTACAAGGAGTCACCTCTTGGGGTCTT 537  
|||||  
Db 18 CCTCTGGTTTGTTCGAGAGGACAAATACATTTTACAAGGAGTCACCTCTTGGGGTCTT 77  
|||||  
QY 538 GGCTGTGCACGCCGCCCAATAGCTGTCTATGTCGTGTTTCAAGGTTTGTACTTGG 597  
|||||  
Db 78 GGCTGTGCACGCCGCCCAATAGCCYGGTGTCTATGTTCTGTTTCAAGGTTTGTACTTGG 137  
|||||  
QY 598 ATTGAGGAGTGATGAGAAATAAT 621  
|||||  
Db 138 ATTGAGGAGTGATGAGAAATAAT 161  
|||||

RESULT 15  
US-10-741-601-14344  
; Sequence 14344, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 14344  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-14344

Query Match 14.5%; Score 90; DB 17; Length 201;  
Best Local Similarity 99.3%; Pred. No. 1.1e-39;  
Matches 140; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 69 GTCCCCAAGGCTTCATCCCTACAGGTTCATCTGGGTGCACACCAAGAGTGAATCTCGA 128  
|||||  
Db 38 GTCCCCAAGGCTTCATCCCTACAGGTTCATCTGGGTGCACACCAAGAGTGAATCTCGA 97  
|||||  
QY 129 ACCGATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGAAAAGATAT 188  
|||||  
Db 98 ACCRCATGTTTCAGAAATAGAGTGTCTAGGCTGTTCTTGGAGCCACACGAAAAGATAT 157  
|||||  
QY 189 TGCCTTGCTAAAGCTAAGCAG 209  
|||||

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 27, 2004, 08:35:06 ; Search time 25 Seconds  
(without alignments)  
4780.049 Million cell updates/sec

Title: US-09-992-095B-53\_COPY\_1044\_1664

Perfect score: 207

Sequence: 1 atgcactctctggaggcac.....aggagtgatgagaaataat 621

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Word size: 1

Total number of hits satisfying chosen parameters: 210334

Minimum DB seq length: 0

Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DBV=xlh  
-Q=cgn2\_1/USPTO.spool/US0992095/runat\_26102004\_085835\_10722/app.query.fasta\_1.775  
-DB=PIR\_79 -OPMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pt  
-NORM=ext -HBASESIZE=500 -MINLEN=0 -MAXLEN=207  
-USER=US0992095 @CGN\_1\_1\_25/runat\_26102004\_085835\_10722 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR\_79.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	5.8	126	A23473	chymotrypsin-like
2	10	4.8	30	A32946	trypsin-like serin
3	10	4.8	61	PS0049	serine proteinase
4	10	4.8	149	S35208	serine proteinase
5	9	4.3	73	S44462	elastase (EC 3.4.2
6	9	4.3	75	A37002	catroxinobin I (EC 3
7	9	4.3	191	S54115	complement factor
8	9	4.3	196	T08808	hypothetical prote
9	8	3.9	30	C32946	serine proteinase
10	8	3.9	43	A61168	cocoanase (EC 3.4.
11	8	3.9	81	A18966	tissue kallikrein
12	8	3.9	85	S44461	elastase (EC 3.4.2
13	8	3.9	90	JE0210	proteinase (EC 3.4
14	8	3.9	94	PC2013	tissue kallikrein

15	8	3.9	161	2	I62744	coagulation factor
16	8	3.9	161	2	I48158	coagulation factor
c 17	8	4.0	173	2	T27650	hypothetical prote
18	7	3.4	25	2	A24807	cytotoxic T-lympho
19	7	3.4	30	2	A61333	trypsin (EC 3.4.21
20	7	3.4	40	2	A49081	capillary permeabi
21	7	3.4	66	2	I52972	kallikrein - mouse
22	7	3.4	96	2	A05308	tissue kallikrein
23	7	3.4	97	2	AF1451	transcription regu
24	7	3.4	99	2	E70679	probable PE protei
25	7	3.4	104	2	S15395	tissue kallikrein-
c 26	7	3.5	119	2	T17003	dormancy-associate
27	7	3.4	126	2	H90505	protein synthesis
28	7	3.4	137	2	S53364	serine proteinase
c 29	7	3.5	146	2	H75394	conserved hypothet
30	7	3.4	149	1	KQMSM	tissue kallikrein
31	7	3.4	151	2	S35205	proteinase 5 - buf
32	7	3.4	152	2	S35209	serine proteinase
33	7	3.4	152	2	S35206	serine proteinase
34	7	3.4	152	2	S35203	serine proteinase
35	7	3.4	156	2	B23863	tissue kallikrein
36	7	3.4	158	2	S35201	serine proteinase
37	7	3.4	158	2	B75422	hypothetical prote
c 38	7	3.5	162	2	F82084	probable 2-demethy
39	7	3.4	165	2	F87661	hypothetical prote
c 40	7	3.5	170	2	H71063	hypothetical prote
41	7	3.4	171	2	B69345	hypothetical prote
42	7	3.4	177	2	S23505	chymase (EC 3.4.21
43	7	3.4	185	2	I40478	conserved hypothet
44	7	3.4	188	2	B32340	tissue kallikrein
45	7	3.4	190	2	D75006	hypothetical prote

ALIGNMENTS

RESULT 1

A23473  
N:Altmann-like proteinase (EC 3.4.21.-) - pig (tentative sequence) (fragments)  
C:Species: Sus scrofa domestica (domestic pig)  
C;Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
C;Accession: A23473  
R;Vered, M.; Gertler, A.; Burstein, Y.  
Int. J. Pept. Protein Res. 27, 183-190, 1986  
A;Reference number: A23473; MUID:86194934; PMID:3634756  
A;Accession: A23473  
A;Molecule type: protein  
A;Residues: 1-126 <VER>  
A;Cross-references: UNIPROT:Q7M325  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: hydrolase; serine proteinase

Alignment Scores:  
Pred. No.: 0.00589 Length: 126  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.80% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x A23473 (1-126)

QY 454 AGTTGCCAGGTGACAGTGGAGTCTCTGTGTTGC 489

Db 95 SerCysGlnGlyAspSerGlyGlyProLeuValCys 106

RESULT 2

A32946

trypsin-like serine proteinase (EC 3.4.21.-) 1 - nematode (Anisakis simplex) (fragments)  
C;Species: Anisakis simplex  
C;Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 20-Sep-1999  
C;Accession: A32946  
R;Sakanari, J.A.; Staunton, C.E.; Eakin, A.E.; Craik, C.S.; McKerrow, J.H.

Proc. Natl. Acad. Sci. U.S.A. 86, 4863-4867, 1989

A:Title: Serine proteases from nematode and protozoan parasites: isolation of sequence h  
A:Reference number: A32946; MUID:89296904; PMID:2662185  
A:Accession: A32946  
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A:Molecule type: DNA  
A:Residues: 1-30 <SAK>  
C:Keywords: hydrolase; serine proteinase

Alignment Scores:  
Pred. No.: 0.653 Length: 30  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.83% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x A32946 (1-30)

QY 451 GACAGTCCAGGTGACAGTGGAGTCT 480

Db 21 AspSerCysGlnGlyAspSerGlyGlyPro 30

RESULT 3

PS0049 serine proteinase (EC 3.4.21.-) 3 - fruit fly (*Drosophila melanogaster*) (fragment)

C:Species: *Drosophila melanogaster*  
C>Date: 07-Jun-1990 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
C:Accession: PS0049  
R:Yun, Y.; Davis, R.L.  
Mol. Cell. Biol. 9, 692-700, 1989  
A:Title: Levels of RNA from a family of putative serine protease genes are reduced in D  
A:Reference number: JS0260; MUID:89219063; PMID:2469005  
A:Accession: PS0049  
A:Molecule type: DNA  
A:Residues: 1-61 <YUN>  
A:Cross-references: UNIPROT:P17207; GB:M24380; NID:gl58409; PIDN:AAB02551.1; PID:gl58412  
A:Experimental source: strain dnc mutant  
C:Genetics:  
A:Gene: SER3

A:Cross-references: FlyBase:FBgn0003358

A:Map position: 3 99C-D

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-53/Domain: trypsin homology (fragment) <TRY>

F:30/Binding site: substrate (Val) #status predicted

Alignment Scores:  
Pred. No.: 0.573 Length: 61  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.83% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x PS0049 (1-61)

QY 457 TCCAGGTGACAGTGGAGTCTCTGTT 486

Db 7 CysGlnGlyAspSerGlyGlyProLeuVal 16

RESULT 4

S35208 serine proteinase (EC 3.4.21.-) 8 - buffalo fly (fragment)

C:Species: *Haematobia irritans exigua* (buffalo fly)  
C>Date: 10-Dec-1993 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C:Accession: S35208; S42696  
R:Elvin, C.M.; Whan, V.; Riddles, P.W.  
Mol. Gen. Genet. 240, 132-139, 1993  
A:Title: A family of serine protease genes expressed in adult buffalo fly (*Haematobia ir*  
A:Reference number: S35201; MUID:93341451; PMID:8341258  
A:Accession: S35208  
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-149 <ELV>

A:Cross-references: UNIPROT:Q06784; EMBL:Z22567

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-149/Domain: trypsin homology (fragment) <TRY>

F:7,52,146/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
Pred. No.: 0.485 Length: 149  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.83% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x S35208 (1-149)

QY 451 GACAGTTCAGGTGACAGTGGAGTCT 480

Db 140 AspSerCysGlnGlyAspSerGlyGlyPro 149

RESULT 5

S44462

elastase (EC 3.4.21.-) 2B - horse (fragments)

C:Species: *Equus caballus* (domestic horse)

C>Date: 20-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 09-Jul-2004

C:Accession: S44462

R:Dubin, A.; Potempa, J.; Travis, J.

Biochem. J. 300, 401-406, 1994

A:Title: Structural and functional characterization of elastases from horse neutrophils.

A:Reference number: S44461; MUID:94271153; PMID:7516152

A:Accession: S44462

A:Molecule type: protein

A:Residues: 1-31;32-56;57-73 <DUB>

A:Cross-references: UNIPROT:P37358

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-73/Domain: trypsin homology (fragments) <TRY>

Alignment Scores:  
Pred. No.: 5.11 Length: 73  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.35% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x S44462 (1-73)

QY 463 GGTGACAGTGGAGTCTCTGTTTC 489

Db 62 GlyAspSerGlyGlyProLeuValCys 70

RESULT 6

A37002

catroxin I (EC 3.4.21.-) - western diamondback rattlesnake (fragments)

C:Species: *Crotalus atrox* (western diamondback rattlesnake)

C>Date: 14-Feb-1992 #sequence\_revision 14-Feb-1992 #text\_change 09-Jul-2004

C:Accession: A37002

R:Pirkle, H.; Theodor, I.; Lopez, R.

Thromb. Res. 56, 159-168, 1989

A:Title: Catroxin, a weakly thrombin-like enzyme from the venom of *Crotalus atrox*.

A:Reference number: A37002; MUID:90141479; PMID:2617466

A:Accession: A37002

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-75 <PIR>

A:Cross-references: UNIPROT:Q7LZF5

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase; venom

F:1-75/Domain: trypsin homology (fragments) <TRY>



Alignment Scores:  
Pred. No.: 5.08 Length: 75  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.35% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x A37002 (1-75)

QY 34 GAGTGGGTGACTGCTGCTGCCACTGC 60

Db 34 GluTrpValLeuThrAlaAlaHisCys 42

RESULT 7

S54115

complement factor D (EC 3.4.21.46) - pig (fragment)

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004

C;Accession: S54115

R;Nicolas, N.

submitted to the EMBL Data Library, April 1995

A;Reference number: S54115

A;Accession: S54115

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-191 <NIC>

A;Cross-references: UNIPROT:P51779; EMBL:Z49058; MID:g773264; PIDN:CAA88844.1; PID:g7732

C;Superfamily: trypsin; trypsin homology

C;Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase

F;1-181/Domain: trypsin homology (fragment) <TRY>

Alignment Scores:  
Pred. No.: 4.27 Length: 191  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.35% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x S54115 (1-191)

QY 463 GGTGACAGTGGAGTCTCTGTTGTC 489

Db 139 GlyAspSerGlyGlyProLeuValCys 147

RESULT 8

T08808

hypothetical protein DKFP586J1923.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004

C;Accession: T08808

R;Ansoorge, W.; Wirkner, U.; Mewes, H.W.; Gaasenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A;Reference number: Z16472

A;Accession: T08808

A;Molecule type: mRNA

A;Residues: 1-196 <ANS>

A;Cross-references: UNIPROT:Q9UKR3; EMBL:AL050220

A;Experimental source: adult uterus; clone DKFP586J1923

C;Genetics:

A;Note: DKFP586J1923.1

C;Superfamily: trypsin; trypsin homology

Alignment Scores:  
Pred. No.: 4.25 Length: 196  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.35% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x T08808 (1-196)

QY 37 TGGGTGTTGACTGCTGCCACTGCTTG 63

Db 86 TrpValLeuThrAlaAlaHisCysLeu 94

RESULT 9

C32946

serine proteinase (EC 3.4.21.-) 3 - nematode (Anisakis simplex) (fragments)

C;Species: Anisakis simplex

C;Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 20-Sep-1999

C;Accession: C32946

R;Sakanari, J.A.; Staunton, C.E.; Eakin, A.E.; Craik, C.S.; McKerrow, J.H.

Proc. Natl. Acad. Sci. U.S.A. 86, 4863-4867, 1989

A;Title: Serine proteases from nematode and protozoan parasites: isolation of sequence

A;Reference number: A32946; MUID:89296904; PMID:2662185

A;Accession: C32946

A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tr

A;Molecule type: DNA

A;Residues: 1-30 <SAK>

C;Keywords: hydrolase; serine proteinase

Alignment Scores:  
Pred. No.: 55.5 Length: 30  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.86% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x C32946 (1-30)

QY 457 TGCCAGGTGACAGTGGAGTCTCT 480

Db 23 CysGlnGlyAspSerGlyGlyPro 30

RESULT 10

A61168

cocoonase (EC 3.4.21.-) precursor - polyphemus moth (fragments)

C;Species: Antheraea polyphemus (polyphemus moth)

C;Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 09-Jul-2004

C;Accession: A61168

R;Kramer, K.J.; Felested, R.L.; Law, J.H.

J. Biol. Chem. 248, 3021-3028, 1973

A;Title: Cocoonase. V. Structural studies on an insect serine protease.

A;Reference number: A61168; MUID:73166540; PMID:4735570

A;Accession: A61168

A;Molecule type: protein

A;Residues: 1-43 <KRA>

A;Cross-references: UNIPROT:Q7M3M4

C;Keywords: hydrolase; serine proteinase; zymogen

F;1-13/Domain: activation peptide #status predicted <APT>

F;14-43/Product: cocoonase (fragment) #status experimental <MAT>

Alignment Scores:  
Pred. No.: 51.9 Length: 43  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.86% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x A61168 (1-43)

QY 457 TGCCAGGTGACAGTGGAGTCTCT 480

Db 30 CysGlnGlyAspSerGlyGlyPro 37

RESULT 11

A18966

tissue kallikrein (EC 3.4.21.35) mGK-22 - mouse (fragments)

N;Alternate names: epidermal growth factor-binding protein type A; nerve growth factor

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1988 #sequence\_revision 06-Feb-1995 #text\_change 18-Jul-1997

C;Accession: A18966; A38356; C18966  
 R;Anundi, H.; Ronne, H.; Peterson, P.A.; Rask, L.  
 Eur. J. Biochem. 129, 365-371, 1982  
 A;Title: Partial amino-acid sequence of the epidermal growth-factor-binding protein.  
 A;Reference number: A91126; MUID:83105150; PMID:6295764  
 A;Accession: A18966  
 A;Molecule type: protein  
 A;Residues: 1-81 <AAU>  
 R;Fahnestock, M.; Woo, J.E.; Lopez, G.A.; Snow, J.; Walz, D.A.; Arici, M.J.; Mobley, W.C.  
 Biochemistry 30, 3443-3450, 1991  
 A;Title: Beta-NGF-endopeptidase: structure and activity of a kallikrein encoded by the g  
 A;Reference number: A38356; MUID:91190897; PMID:2012805  
 A;Accession: A38356  
 A;Molecule type: protein  
 A;Residues: 1-30 <FAH>  
 C;Superfamily: trypsin; trypsin homology  
 C;Keywords: hydrolase; serine proteinase  
 F;1-81/Domain: trypsin homology (fragments) <TRY>  
 Alignment Scores:  
 Pred. No.: 46.1 Length: 81  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.86% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x A18966 (1-81)  
 QY 37 TGGGTGTGACTGCTGCCACTGC 60  
 DB 35 TrpValLeuThrAlaAlaHisCys 42  
 RESULT 12  
 S44461  
 elastase (EC 3.4.21.-) 2A - horse (fragments)  
 C;Species: Equus caballus (domestic horse)  
 C;Date: 20-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 09-Jul-2004  
 C;Accession: S44461  
 R;Dubin, A.; Potempa, J.; Travis, J.  
 Biochem. J. 300, 401-406, 1994  
 A;Title: Structural and functional characterization of elastases from horse neutrophils.  
 A;Reference number: S44461; MUID:94271153; PMID:7516152  
 A;Accession: S44461  
 A;Molecule type: protein  
 A;Residues: 1-34;35-59;60-85 <DUB>  
 A;Cross-references: UNIPROT:P17357  
 C;Superfamily: trypsin; trypsin homology  
 C;Keywords: hydrolase; serine proteinase  
 Alignment Scores:  
 Pred. No.: 45.7 Length: 85  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.86% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x S44461 (1-85)  
 QY 4 CACTTCTGTGAGGACCTTGATA 27  
 DB 27 HisPheCysGlyGlyThrLeuile 34  
 RESULT 13  
 J50210  
 proteinase (EC 3.4.21.-) serine-like, NES1 - human  
 C;Species: Homo sapiens (man)  
 C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 07-May-1999  
 C;Accession: J50210  
 R;Luo, L.; Herbrick, J.A.; Scherer, S.W.; Beatty, B.; Squire, J.; Diamandis, E.P.  
 Biochem. Biophys. Res. Commun. 247, 580-586, 1998  
 A;Title: Structure characterization and mapping of the normal epithelial cell-specific

A;Reference number: JE0210; MUID:98321170; PMID:9647736  
 A;Accession: JE0210  
 A;Molecule type: mRNA  
 A;Residues: 1-90 <LUO>  
 A;Cross-references: GB:AF055481  
 C;Genetics:  
 A;Gene: GDB:PRSSL1; NES1  
 A;Map position: 19q13.3-13.4  
 C;Keywords: hydrolase  
 Alignment Scores:  
 Pred. No.: 45.3 Length: 90  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.86% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x JE0210 (1-90)  
 QY 466 GACAGTGGAGTCTCTGTTTC 489  
 DB 74 AspSerGlyGlyProLeuValCys 81  
 RESULT 14  
 PC2013  
 tissue kallikrein (EC 3.4.21.35) mK1, submandibular - mouse (fragments)  
 N;Alternate names: proteinase F  
 C;Species: Mus musculus (house mouse)  
 C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 09-Jul-2004  
 C;Accession: PC2013  
 R;Hosoi, K.; Tsunawawa, S.; Kurihara, K.; Aoyama, H.; Ueha, T.; Sakiyama, F.  
 J. Biochem. 115, 137-143, 1994  
 A;Title: Identification of mK1, a true tissue (glandular) kallikrein of mouse submandib  
 s.  
 A;Reference number: PC2013; MUID:94245648; PMID:8188620  
 A;Accession: PC2013  
 A;Molecule type: protein  
 A;Residues: 1-94 <HOS>  
 A;Cross-references: UNIPROT:Q7M0B4  
 A;Experimental source: submandibular gland  
 C;Genetics:  
 A;Gene: mK1  
 C;Superfamily: trypsin; trypsin homology  
 C;Keywords: hydrolase; serine proteinase; submandibular gland  
 F;1-94/Domain: trypsin homology (fragments) <TRY>  
 Alignment Scores:  
 Pred. No.: 44.9 Length: 94  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.86% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x PC2013 (1-94)  
 QY 37 TGGGTGTGACTGCTGCCACTGC 60  
 DB 35 TrpValLeuThrAlaAlaHisCys 42  
 RESULT 15  
 I62744  
 coagulation factor Xa (EC 3.4.21.6) - rhesus macaque (fragment)  
 C;Species: Macaca mulatta (rhesus macaque)  
 C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 09-Jul-2004  
 C;Accession: I62744  
 R;Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.  
 Eur. J. Haematol. 52, 162-168, 1994  
 A;Title: Analysis of the partial nucleotide sequences and deduced primary structures of  
 A;Reference number: I46196; MUID:94222160; PMID:8168596  
 A;Accession: I62744  
 A;Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA  
A:Residues: 1-161 <RES>  
A:Cross-references: UNIPROT:Q28511; GB:D21214; NID:G415307; PIDN:BAA04755.1; PID:G455395  
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglutamate  
F:1-161/domain: trypsin homology (fragment) <TRY>

Alignment Scores:                      40.6                      Length:                      161  
Pred. No.:                      8.00                      Matches:                      8  
Score:                      100.00%                      Conservative:                      0  
Percent Similarity:                      100.00%                      Mismatches:                      0  
Best Local Similarity:                      3.86%                      Indels:                      0  
Query Match:                      2                      Gaps:                      0  
DB:                      2

US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x 162744 (1-161)

QY                      457 TGCCAGGGTGACAGTGGAGTCCT 480  
                         |||||  
DB                      119 CysGlnGlyAspSerGlyGlyPro 126  
                         |||||

Search completed: October 27, 2004, 09:12:34  
Job time : 29 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: October 27, 2004, 01:56:25 ; Search time 2779 Seconds  
(without alignments)  
8142.884 Million cell updates/sec

Title: US-09-992-095B-53\_COPY\_1044\_1664  
Perfect score: 621  
Sequence: 1 atgcactctgtggaggcac.....aggagtgatgagaataat 621

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size : 0

Total number of hits satisfying chosen parameters: 41269900

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	404	65.1	523	6	CB164684 K-EST0225
2	348	56.0	616	5	BU073846 in25c11.y
3	342	55.1	433	7	H90220 yu85g08.r1
4	318	51.2	564	6	CB161593 K-EST0221
5	315	50.7	475	7	N91337 za14h05.r1
6	314	50.6	543	6	CB161910 K-EST0221
7	313	50.4	517	6	CB162370 K-EST0222
8	309	49.8	481	7	H60805 yr45g09.r1
9	293	47.5	575	7	T84554 yd53d03.r1
10	293	47.2	381	7	H37877 yp57f06.r1
11	287	46.2	504	7	N77239 yv44d01.r1
12	287	46.2	611	1	AV662084 AV662084
13	286	46.1	541	1	AA722885 z89g06.s
14	278	44.8	593	2	BE348267 hw21e11.x
15	272	43.8	594	1	AV661991 AV661991
16	268	43.2	459	7	R94305 yq41c07.r1
17	261	42.0	303	7	T73187 yc69h06.r1
18	248	39.9	424	7	H73861 yel12a02.r1
19	228	36.7	466	1	AI948806 wg26e12.x
20	219	35.3	595	2	AW950594 EST162664
21	214	34.5	563	2	AW271976 xs12d10.x
22	203	32.7	341	7	T70507 yc38g05.r1
23	198	31.9	614	1	AV693554 AV693554
24	196	31.6	553	1	AI633979 wa32c10.x

C 25	193	31.1	621	1	AI377474	AI377474	tc37a03.x
C 26	190	30.6	444	1	AI640315	AI640315	wa15g08.x
C 27	189	30.4	405	2	AW950595	EST162665	AW950595
C 28	189	30.4	589	7	H73871	YB14d05.r1	H73871
C 29	186	30.0	464	2	BE326689	hr63d09.x	BE326689
C 30	185	29.8	444	7	T73867	yc56g11.r1	T73867
C 31	182	29.3	526	6	CD608134	56072712H	CD608134
C 32	182	29.3	526	6	CD608136	56072820H	CD608136
C 33	181	29.1	360	1	AV654539	AV654539	AV654539
C 34	181	29.1	607	1	AV662061	AV662061	AV662061
C 35	179	28.8	457	1	AI631756	wa35g01.x	AI631756
C 36	178	28.7	368	7	H95494	YV13f06.r1	H95494
C 37	178	28.7	558	2	AW104579	xd76f03.x	AW104579
C 38	169	27.2	545	1	AV720081	AV720081	AV720081
C 39	163	26.2	416	7	T68303	YC41d12.s1	T68303
C 40	147	23.7	462	1	AI768449	wh22a07.x	AI768449
C 41	144	23.2	393	7	T51771	YB27h05.r1	T51771
C 42	144	23.2	398	1	AA953491	on72c03.s	AA953491
C 43	144	23.2	423	4	BG565038	60253832	BG565038
C 44	144	23.2	430	2	BF197584	7085506.x	BF197584
C 45	144	23.2	462	1	AA995929	ou93h04.s	AA995929

ALIGNMENTS

RESULT 1  
CB164684  
LOCUS  
DEFINITION  
K-EST0225947 L17N670205n1 Homo sapiens cDNA clone  
L17N670205n1-43-C07 5', mRNA sequence.  
523 bp mRNA linear EST 30-JAN-2003  
ACCESSION  
CB164684  
VERSION  
CB164684.1 GI:28150810  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
REFERENCE  
1 (bases 1 to 523)  
AUTHORS  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE  
21C Frontier Korean EST Project 2001  
JOURNAL  
Unpublished (2002)  
COMMENT  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongseung@mail.kribb.re.kr  
Plate: 43 row: C column: 07  
High quality sequence stop: 523.  
Location/Qualifiers  
1..523  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L17N670205n1-43-C07"  
/sex="F"  
/lab\_host="Top10P"  
/clone\_lib="L17N670205n1"  
/notes="Organ: Liver; Vector: pTVT3-Pac; Site 1: ECORI;  
Site 2: NotI; The library was contributed by the Soares  
Laboratory and it was constructed as described by Bonaldo,  
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research  
6(9): 791-806. RNA was prepared from harvested cell  
culture."

ORIGIN

Query Match 65.1%; Score 404; DB 6; Length 523;  
Best Local Similarity 99.8%; Pred. No. 4.6e-208;  
Matches 454; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 167 TGGAGCCACACAGAAAGATATTGCTTGTCTAAGCTAAGAGTCTCTCCCTCATCTG 226  
 Db 31 TGGAGCCACACAGAAAGATATTGCTTGTCTAAGCTAAGAGTCTCTCCCTCATCTG 90  
 QY 227 ACAAGATTAATCCAGCTTGTCTGCGATCCCAATATATGCTGCTGACCGGACCGAAT 286  
 Db 91 ACAAGATTAATCCAGCTTGTCTGCGATCCCAATATATGCTGCTGACCGGACCGAAT 150  
 QY 287 GTTTCATCACTGGCTGGGAGAAACCCAAAGTACTTTTGGAGCTGGCTTCTCAAGGAAG 346  
 Db 151 GTTTCATCACTGGCTGGGAGAAACCCAAAGTACTTTTGGAGCTGGCTTCTCAAGGAAG 210  
 QY 347 CCCAGCTCCCTGTGATTGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAG 406  
 Db 211 CCCAGCTCCCTGTGATTGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAATGGAAG 270  
 QY 407 TCCATCCACCGAATCTGTCTGGGCAATTTGGCCGAGGAGTCTGACAGTTTGCAGGGTG 466  
 Db 271 TCCATCCACCGAATCTGTCTGGGCAATTTGGCCGAGGAGTCTGACAGTTTGCAGGGTG 330  
 QY 467 ACAGTGGAGGCTCTCTGTTTCTTCGAGAGGACAAATACATTTTCAAGGAGTCACTT 526  
 Db 331 ACAGTGGAGGCTCTCTGTTTCTTCGAGAGGACAAATACATTTTCAAGGAGTCACTT 390  
 QY 527 CTTGGGCTTGGCTGTGACGCCCAATAGCCGTGTCTATGTTTCTGTTTCAAGGT 586  
 Db 391 CTTGGGCTTGGCTGTGACGCCCAATAGCCGTGTCTATGTTTCTGTTTCAAGGT 450  
 QY 587 TTGTTACTTGGATTGAGGAGTGATGAGAAATAAT 621  
 Db 451 TTGTTACTTGGATTGAGGAGTGATGAGAAATAAT 485

RESULT 2  
 BU073846  
 LOCUS  
 DEFINITION in25c11.y1 Human Fetal Pancreas 1B Homo sapiens cDNA clone IMAGE: 5', similar to SW:FLMN\_HUMAN P00747 PLASMINOGEN PRECURSOR ;, mRNA sequence.

ACCESSION BU073846  
 VERSION BU073846.1 GI:22515035  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

REFERENCE  
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Teagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.

TITLE Endocrine Pancreas Consortium  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Hiroshi Inoue (hinoue@imgate.wustl.edu)  
 Putative full length read  
 vector to vector length is 617  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 470.  
 Location/Qualifiers

source 1..616  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:"  
 /issue\_type="Fetal Pancreas (4 Pooled Donors, 18 - 20 weeks, Stratagene #738023)"  
 /dev\_stage="Fetal Pancreas"  
 /clone\_lib="Human Fetal Pancreas 1B"  
 /notes="Vector: pBluescript SK(-); Site 1: NotI; Site 2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permutt Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110 E-mail: hinoue@imgate.wustl.edu Tel: 314-362-1916, Fax:314-747-2692."

ORIGIN

Query Match 56.0%; Score 348; DB 5; Length 616;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e-177;  
 Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 274 GACCGGACCGAATGTTTCATCACTGGCTGGGGAGAAACCCAAAGTACTTTTGGAGCTGGC 333  
 Db 3 GACCGGACCGAATGTTTCATCACTGGCTGGGGAGAAACCCAAAGTACTTTTGGAGCTGGC 62  
 QY 334 CTTCTCAAGGAAGCCAGCTCCCTGTGATTGAGAAATAAAGTGTGCAATCGCTATGAGTTT 393  
 Db 63 CTTCTCAAGGAAGCCAGCTCCCTGTGATTGAGAAATAAAGTGTGCAATCGCTATGAGTTT 122  
 QY 394 CTGAATCGAAGAGTCCCAATCCACCGAATCTGTCTGGGCAATTTGGCGGAGGCACTGAC 453  
 Db 123 CTGAATCGAAGAGTCCCAATCCACCGAATCTGTCTGGGCAATTTGGCGGAGGCACTGAC 182  
 QY 454 AGTTGCCAGGCTGACAGTGGAGGTCCTCTGTTTGGCTTCGAGAGGACAAATACATTTTA 513  
 Db 183 AGTTGCCAGGCTGACAGTGGAGGTCCTCTGTTTGGCTTCGAGAGGACAAATACATTTTA 242  
 QY 514 CAAGGAGTCACTTCTTGGGGTCTTGGTGTGTCACGCCCAATTAAGCTGTGTCTATGTTT 573  
 Db 243 CAAGGAGTCACTTCTTGGGGTCTTGGTGTGTCACGCCCAATTAAGCTGTGTCTATGTTT 302  
 QY 574 CGTGTTCAGAGTTTGTACTTGGATTGAGGAGTGTGAGAAATAAT 621  
 Db 303 CGTGTTCAGAGTTTGTACTTGGATTGAGGAGTGTGAGAAATAAT 350

RESULT 3

H90220

LOCUS

DEFINITION H90220 yu85908.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:240638 5' similar to gb:X05199 PLASMINOGEN PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION H90220

VERSION H90220.1 GI:1080650

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 475)

# REFERENCE AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Ruffing, T., Soares, M., Tan, F., Trevaaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

The WashU-Merck EST Project

# TITLE JOURNAL

Unpublished (1995)

# COMMENT

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Seq primer: mob.REGA+ET

High quality sequence stop: 296.

# FEATURES

source

1..475

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="GDB:3800346"

/db\_xref="taxon:9606"

/clone="IMAGE:292569"

/sex="male"

/dev\_stage="20 week-post conception fetus"

/lab\_host="DH10B (ampicillin resistant)"

/clone\_lib="Soares fetal liver spleen 1NFLS"

/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)

with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;

1st strand cDNA was primed with a Pac I - oligo(dT) primer

[5' AACGGAAGTAATTAAGATCTTTTTTTTTTTTTTTT 3']

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified p773 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M.Fatima Bonaldo."

# ORIGIN

Query Match 50.7%; Score 315; DB 7; Length 475;  
Best Local Similarity 100.0%; Pred. No. 1.3e-159; Mismatches 0; Indels 0; Gaps 0;  
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 GTTACTGCTGCCACTGTTGGAGAAGTCTCGAACCGCATGTTCCAGGAATAGAAAGTGTCTAGGCT 101

Db 20 GTTACTGCTGCCACTGTTGGAGAAGTCTCGAACCGCATGTTCCAGGAATAGAAAGTGTCTAGGCT 79

Qy 102 GGGTGCACACCAAGAAGTGAATCTCGAACCGCATGTTCCAGGAATAGAAAGTGTCTAGGCT 161

Db 80 GGGTGCACACCAAGAAGTGAATCTCGAACCGCATGTTCCAGGAATAGAAAGTGTCTAGGCT 139

Qy 162 GTTCTTGGAGCCACACCAAGAAGTGAATCTCGAACCGCATGTTCCAGGAATAGAAAGTGTCTAGGCT 221

Db 140 GTTCTTGGAGCCACACCAAGAAGTGAATCTCGAACCGCATGTTCCAGGAATAGAAAGTGTCTAGGCT 199

Qy 222 CACTGACAAAGTAATCTCCAGCTTGTCTGCCATCCCCAAATATGTGGTCTGACCGGAC 281

Db 200 CACTGACAAAGTAATCTCCAGCTTGTCTGCCATCCCCAAATATGTGGTCTGACCGGAC 259

Qy 282 CGAATGTTTCATCATGGCTGGGAGAAACCAAGGATCTTTTGGAGCTGGCTTCTCAA 341

Db 260 CGAATGTTTCATCATGGCTGGGAGAAACCAAGGATCTTTTGGAGCTGGCTTCTCAA 319

Qy 342 GGAAGCCAGTCCC 356

Db 320 GGAAGCCAGTCCC 334

RESULT 6  
CB161910

LOCUS CB161910 543 bp mRNA linear EST 30-JAN-2003  
DEFINITION K-EST0221967 L17N670205n1 Homo sapiens cDNA clone  
L17N670205n1-41-G11 5', mRNA sequence.

ACCSSION CB161910

VERSION CB161910.1 GI:28148036

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 543)

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,

Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and

Kim, Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongseung@mail.kribb.re.kr

Plate: 41 row: G column: 11

High quality sequence stop: 543.

Location/Qualifiers

1..543

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="L17N670205n1-41-G11"

/sex="p"

/lab\_host="Top10P"

/clone\_lib="L17N670205n1"

/note="Organ: Liver; Vector: p773-Pac; Site 1: EcoRI;

Site 2: NotI; The library was contributed by the Soares

laboratory and it was constructed as described by Bonaldo,

M.F., Lennon, G. and Soares, M.B. (1996), Genome Research

6(9): 791-806. RNA was prepared from harvested cell

culture."

# ORIGIN

Query Match 50.6%; Score 314; DB 6; Length 543;  
Best Local Similarity 100.0%; Pred. No. 4.5e-159; Mismatches 0; Indels 0; Gaps 0;  
Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCACCTTCTGTGGAGCACCTTGATATCCCAAGTGGGTGTGACTGTGCTGCCACTGC 60

Db 230 ATGCACCTTCTGTGGAGCACCTTGATATCCCAAGTGGGTGTGACTGTGCTGCCACTGC 289

Qy 61 TTGGAGAAGTCCCCAAGCGCTTCATCTCCTACAGGTCTCTGGGTGCACACCAAGAGTG 120

Db 290 TTGGAGAAGTCCCCAAGCGCTTCATCTCCTACAGGTCTCTGGGTGCACACCAAGAGTG 349

Qy 121 AATCTCGAACCGCATGTTTCAGGAAATAGAAAGTGTCTAGGCTGTTCTTGGAGCCACACGA 180

Db 350 AATCTCGAACCGCATGTTTCAGGAAATAGAAAGTGTCTAGGCTGTTCTTGGAGCCACACGA 409

Qy 181 AAAGATATTCCTTGTCTTAAAGCTAAGCAGTCTCCCGTCTACCTGACAAAGTAAATCCCA 240

Db 410 AAAGATATTCCTTGTCTTAAAGCTAAGCAGTCTCCCGTCTACCTGACAAAGTAAATCCCA 469

Qy 241 GCTTGTCTGCCATCCCCAAATATGCTGCTGACCGGACCGGAATGTTTCATCACTGGC 300

Db 470 GCTTGTCTGCCATCCCCAAATATGCTGCTGACCGGACCGGAATGTTTCATCACTGGC 529

Qy 301 TGGGGAGAAACCCA 314

Db 530 TGGGGAGAAACCCA 543

RESULT 7  
CB162370



LOCUS	CB162370	517 bp	mRNA	linear	EST 30-JAN-2003
DEFINITION	K-EST0222936 L17N670205n1 Homo sapiens cDNA clone				
	L17N670205n1-42-H08 5', mRNA sequence.				
ACCESSION	CB162370				
VERSION	CB162370.1	GI:28148496			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 517) Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.				
TITLE	21C Frontier Korean EST Project 2001				
JOURNAL	Unpublished (2002)				
COMMENT	Genome Research Center Contact: Kim YS Korea Research Institute of Bioscience & Biotechnology 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.kribb.re.kr Plate: 42 row: H column: 08 High quality sequence stop: 517.				
FEATURES	Location/Qualifiers				
source	1..517				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="L17N670205n1-42-H08"				
	/sex="F"				
	/lab_host="Top10F"				
	/clone_lib="L17N670205n1"				
	/note="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI; The library was contributed by the Soares laboratory and it was constructed as described by Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome Research 6(9): 791-806. RNA was prepared from harvested cell culture."				
ORIGIN					
Query Match	50.4%; Score 313; DB 6; Length 517;				
Best Local Similarity	100.0%; Pred. No. 1.6e-158;				
Matches 313; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGCACCTTCGTGGAGGCACCTTGATATCCCGAGAGTGGGTGTTGACTGCTGCCACTGC	60		
Db	205	ATGCACCTTCGTGGAGGCACCTTGATATCCCGAGAGTGGGTGTTGACTGCTGCCACTGC	264		
Qy	61	TTGGAGAGTCCCGAAGGCTTCATCCTCAAGAGTCACTCCTGGGTGCACACCAAGAAGTG	120		
Db	265	TTGGAGAGTCCCGAAGGCTTCATCCTCAAGAGTCACTCCTGGGTGCACACCAAGAAGTG	324		
Qy	121	AATCTCGAACCGCATGTTTCAGGAATAGAAAGTGTCTAGGCTGTTCTTGAGGCCACGACA	180		
Db	325	AATCTCGAACCGCATGTTTCAGGAATAGAAAGTGTCTAGGCTGTTCTTGAGGCCACGACA	384		
Qy	181	AAAGATATTGCCCTTCTTAAGCTTACGACGTCTCGCGTCACTACGACAAAGTAATCCCA	240		
Db	385	AAAGATATTGCCCTTCTTAAGCTTACGACGTCTCGCGTCACTACGACAAAGTAATCCCA	444		
Qy	241	GCTTGTCTGCATCCCAAAATATGTGGTGCCTGACCGGACCGAATGTTTTCATCACTGGC	300		
Db	445	GCTTGTCTGCATCCCAAAATATGTGGTGCCTGACCGGACCGAATGTTTTCATCACTGGC	504		
Qy	301	TGGGGAGAAACCC	313		
Db	505	TGGGGAGAAACCC	517		
RESULT 8	H60805				

```

Qy 241 CTTGCTGTCATCCCAAAATATGTCCTGACCGGACCGAATGTTTCATCACTGGC 300
Db 320 CTTGCTGTCATCCCAAAATATGTCCTGACCGGACCGAATGTTTCATCACTGGC 379

Qy 301 TGGGAGAA 309
Db 380 TGGGAGAA 388

RESULT 9
T84554
LOCUS
DEFINITION
IMAGE:111941 5' similar to gb:X05199 PLASMINOGEN PRECURSOR
(HUMAN); mRNA sequence.
ACCESSION
T84554
VERSION
T84554.1 GI:712906
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 575)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, F., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 762
High quality sequence stops: 280
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 762 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 280.
Location/Qualifiers
1..575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:467558"
/db_xref="taxon:9606"
/clone="IMAGE:111941"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACTGGAAGAAATTAATAAGATCTTTTTCATCACTGGGAGCT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

FEATURES
source
1..575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:467558"
/db_xref="taxon:9606"
/clone="IMAGE:111941"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACTGGAAGAAATTAATAAGATCTTTTTCATCACTGGGAGCT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN
Query Match 47.5%; Score 295; DB 7; Length 575;
Best Local Similarity 100.0%; Pred. No. 1e-148;
Matches 295; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 GTGCTAGGCTGTTCTTGAGGCCACACGAAAGATATGCTTAAAGCTAAGCAGT 210
Db 4 GTGCTAGGCTGTTCTTGAGGCCACACGAAAGATATGCTTAAAGCTAAGCAGT 63

```

```

Qy 211 CTTGCGCTCATCTACGACAAAGTAATCCCAAGTGTCTGCCATCCCAAAATATGTCGTC 270
Db 64 CTTGCGCTCATCTACGACAAAGTAATCCCAAGTGTCTGCCATCCCAAAATATGTCGTC 123

Qy 271 GCTGACCGGACCGAATGTTTCATCACTGGTGGGGAGAAACCCCAAGTACTTTTGGAGCT 330
Db 124 GCTGACCGGACCGAATGTTTCATCACTGGTGGGGAGAAACCCCAAGTACTTTTGGAGCT 183

Qy 331 GGCCTTCTCAAGGAGCCGAGCTCCCTGCTGATTTGAGATAAAGTGTGCAATCGCTATCAG 390
Db 184 GGCCTTCTCAAGGAGCCGAGCTCCCTGCTGATTTGAGATAAAGTGTGCAATCGCTATCAG 243

Qy 391 TTTCTGAATGAAGAGTCCCAATCCACCGAACTCTGTCTGGGCACTTTTGGCCGGAG 445
Db 244 TTTCTGAATGAAGAGTCCCAATCCACCGAACTCTGTCTGGGCACTTTTGGCCGGAG 298

RESULT 10
H37877
LOCUS
DEFINITION
IMAGE:191555 5' similar to gb:X05199 PLASMINOGEN PRECURSOR
(HUMAN); mRNA sequence.
ACCESSION
H37877
VERSION
H37877.1 GI:907376
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 381)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, F., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 843
High quality sequence stops: 278
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 843 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 278.
Location/Qualifiers
1..381
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3761344"
/db_xref="taxon:9606"
/clone="IMAGE:191555"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACTGGAAGAAATTAATAAGATCTTTTTCATCACTGGGAGCT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

```

Query Match 47.2%; Score 293; DB 7; Length 381;  
Best Local Similarity 100.0%; Pred. No. 1.2e-147;  
Matches 293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 GTGCTAGAGCTGTTCTGGAGCCACACGAAAGATATTGCTTGTCTAAAGCTAAGCACT 210  
DB 1 GTGCTAGAGCTGTTCTGGAGCCACACGAAAGATATTGCTTGTCTAAAGCTAAGCACT 60

QY 211 CTGCGCGTCATCTACGACAAAGTAATCCAGCTTGTCTGCCATCCCAATATTGTGTC 270  
DB 61 CTGCGCGTCATCTACGACAAAGTAATCCAGCTTGTCTGCCATCCCAATATTGTGTC 120

QY 271 GCTGACCGGACCAAGTTTTCATCACTGGCTGGGAGAAACCAAGCTACTTTTGGAGCT 330  
DB 121 GCTGACCGGACCAAGTTTTCATCACTGGCTGGGAGAAACCAAGCTACTTTTGGAGCT 180

QY 331 GGCTTTCTCAAGGAAGCCAGCTCCCTGTGATTGAGAATAAAGTGTGCAATCGCTATGAG 390  
DB 181 GGCTTTCTCAAGGAAGCCAGCTCCCTGTGATTGAGAATAAAGTGTGCAATCGCTATGAG 240

QY 391 TTTCTGAATGGAAGTCCATCCAGCACTGTGCTGGGCAATTTGGCCGG 443  
DB 241 TTTCTGAATGGAAGTCCATCCAGCACTGTGCTGGGCAATTTGGCCGG 293

RESULT 11  
N77239  
LOCUS  
DEFINITION  
Yv4401.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
IMAGE:245569 5' similar to gb:X05199 PLASMINOGEN PRECURSOR  
(HUMAN); mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 504)  
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
Chissole,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,  
Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,  
Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,  
Roehlfing,T., Schellenberg,K., Soares,M.B., Tan,E., Thierry-Mieg,J.,  
Trevisan,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.  
and Marra,M.

Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
97044478  
8889549  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 832 Std Error: 0.00  
Seq primer: reverse EST  
High quality sequence stop: 479.

FEATURES  
source  
1. 504  
/mol\_type="mRNA"  
/db\_xref="GDB:3794815"  
/db\_xref="taxon:9606"  
/clone="IMAGE:245569"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"

/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGGAAGTAATTAAGATCTTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified p7T3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN  
Query Match 46.2%; Score 287; DB 7; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.2e-144;  
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 AGGCTGTTCTTGGAGCCACACGAAAGATATTGCTTGTCTAAAGCTAAGCACTCTGCC 216  
DB 11 AGGCTGTTCTTGGAGCCACACGAAAGATATTGCTTGTCTAAAGCTAAGCACTCTGCC 70

QY 217 GTCATCACTGACAAAGTAATCCAGCTTGTCTGCCATCCCAATATTATGTGCTGCTGAC 276  
DB 71 GTCATCACTGACAAAGTAATCCAGCTTGTCTGCCATCCCAATATTATGTGCTGCTGAC 130

QY 277 CGGACCGAATGTTTTCATCACTGGCTGGGAGAAACCCAGGTACTTTTCGAGCTGCGCTT 336  
DB 131 CGGACCGAATGTTTTCATCACTGGCTGGGAGAAACCCAGGTACTTTTCGAGCTGCGCTT 190

QY 337 CTCAGGAAGCCAGCTCCCTGTGATTGAGATAAAGTGTCAATCGCTATGAGTTCTG 396  
DB 191 CTCAGGAAGCCAGCTCCCTGTGATTGAGATAAAGTGTCAATCGCTATGAGTTCTG 250

QY 397 AATGGAAGAGTCCAATCCACCACTCTGTCTGGGCAATTTGGCCGG 443  
DB 251 AATGGAAGAGTCCAATCCACCACTCTGTCTGGGCAATTTGGCCGG 297

RESULT 12  
AV662084  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 611)  
Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,  
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,  
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,  
Hu,G., Gu,J., Chen,Z. and Han,Z.  
Insight into hepatocellular carcinogenesis at transcriptome level  
by comparing gene expression profiles of hepatocellular carcinoma  
with those of corresponding noncancerous liver  
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
21625106  
11752456  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1. 611  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="GLCHAG06"  
/tissue\_type="corresponding non cancerous liver tissue"



```
source
1. 593
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:3183596"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/notes="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaudo."
```

ORIGIN

Query Match 44.8%; Score 278; DB 2; Length 593;  
Best Local Similarity 100.0%; Pred. No. 1.8e-139; Indels 0; Gaps 0;  
Matches 278; Conservative 0; Mismatches 0

QY 344 AAGCCAGCTCCCTGTGATTGAGATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAA 403  
DB 505 AAGCCAGCTCCCTGTGATTGAGATAAAGTGTGCAATCGTATGAGTTTCTGAATGGAA 446

QY 404 GAGTCAATCCACCGAATCTGTGCTGGGCAATTTGGCCGAGGACATGACAGTTGCCAGG 463  
DB 445 GAGTCAATCCACCGAATCTGTGCTGGGCAATTTGGCCGAGGACATGACAGTTGCCAGG 386

QY 464 GTGACAGTGGAGTCTCTGTTTGTCTCGAGAAGGACAAATACATTTTACAGGAGTCA 523  
DB 385 GTGACAGTGGAGTCTCTGTTTGTCTCGAGAAGGACAAATACATTTTACAGGAGTCA 326

QY 524 CTTCCTTGGGCTCTTGGCTGTGCACGCCCCCAATAGCCTGTGTCTATGTTTCTGTTTCAA 583  
DB 325 CTTCCTTGGGCTCTTGGCTGTGCACGCCCCCAATAGCCTGTGTCTATGTTTCTGTTTCAA 266

QY 584 GGTGTTGTTACTTGGATTGAGGAGTGTATGAGAAATAAT 621  
DB 265 GGTGTTGTTACTTGGATTGAGGAGTGTATGAGAAATAAT 228

RESULT 15

AV661991  
LOCUS AV661991 GLC Homo sapiens cDNA clone GLCGZG06 3', mRNA sequence. EST 16-JAN-2002  
DEFINITION AV661991  
ACCESSION AV661991  
VERSION AV661991.1 GI:9883005  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 594)  
XU,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,  
Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,  
Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,  
Hu,G., Gu,J., Chen,Z. and Han,Z.  
Insight into hepatocellular carcinogenesis at transcriptome level  
by comparing gene expression profiles of hepatocellular carcinoma  
with those of corresponding noncancerous liver  
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)  
21625106  
11752456  
Contact: Zequang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES

Source

Location/Qualifiers

1. 594

/organism="Homo sapiens"

/mol\_type="mrna"

/db\_xref="taxon:9606"

/clone="GLCGZG06"

/tissue\_type="corresponding non cancerous liver tissue"

/dev\_stage="Adult"

/lab\_host="SOLR"

/clone\_lib="GLC"

/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN

Query Match 43.8%; Score 272; DB 1; Length 594;  
Best Local Similarity 100.0%; Pred. No. 3.3e-136; Indels 0; Gaps 0;  
Matches 272; Conservative 0; Mismatches 0

QY 1 ATGCACCTTCTGTGGAGGACCTTGATATCCCCCAGAGTGGTGTGACTGCTGCCCACTGC 60  
DB 70 ATGCACCTTCTGTGGAGGACCTTGATATCCCCCAGAGTGGTGTGACTGCTGCCCACTGC 129

QY 61 TTGGAGAAGTCCCCAAGGCCTTCATCTACAGGTCTATCTGGTGCACACCAAGAAGTG 120  
DB 130 TTGGAGAAGTCCCCAAGGCCTTCATCTACAGGTCTATCTGGTGCACACCAAGAAGTG 189

QY 121 ATCTCGAACCCGATGTTTCAGGAATAGAAAGTGTCTAGGCTGTCTTGGAGGCCACACGA 180  
DB 190 ATCTCGAACCCGATGTTTCAGGAATAGAAAGTGTCTAGGCTGTCTTGGAGGCCACACGA 249

QY 181 AAAGATATTTGCTTGTCTAAAGCTAAGCAAGTCTCTGCCGTATCATCTGACAAAGTAATCCCA 240  
DB 250 AAAGATATTTGCTTGTCTAAAGCTAAGCAAGTCTCTGCCGTATCATCTGACAAAGTAATCCCA 309

QY 241 GCTTGTCTGCCATCCCCAAATTTATGTGTCGC 272  
DB 310 GCTTGTCTGCCATCCCCAAATTTATGTGTCGC 341

Search completed: October 27, 2004, 07:37:32  
Job time : 2786 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: October 27, 2004, 07:39:21 ; Search time 111.5 Seconds  
(without alignments)  
6409.106 Million cell updates/sec

Title: US-09-992-095b-53\_COPY\_1044\_1664  
Perfect score: 207  
Sequence: 1 agcactctctggaggcac.....aggagtgatgagaaataat 621

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1825181 seqs, 575374646 residues  
Word size: 1  
Total number of hits satisfying chosen parameters: 1566568

Minimum DB seq length: 0  
Maximum DB seq length: 207

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODE=frame+n2p.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US0992095/runat\_26102004\_085835\_10710/app\_query.fasta\_1.775  
-DB=Uniprot\_02 -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pt  
-NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=207  
-USER=US0992095 @CGN 1 1 101 @runat\_26102004\_085835\_10710 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Uniprot\_02.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	18.8	73	2 Q9TV90	Q9TV90 equus caball
2	32	15.5	54	2 Q6JDI3	Q6JDI3 canis famil
3	14	6.8	176	2 Q6SLI2	Q6SLI2 canis famil
4	14	6.8	176	2 AAR19224	Aar19224 canis fam
5	13	6.3	87	2 Q9CQ78	Q9CQ78 m mus muscu
6	13	6.3	176	2 Q1VQ8	Q1VQ8 mus musculu
7	12	5.8	126	2 Q7M325	Q7M325 sus scrofa
8	12	5.8	164	2 Q9DC82	Q9DC82 mus musculu
9	12	5.8	165	1 TRY3 LUCCU	P35043 lucilia cup
10	11	5.3	119	2 Q9NR68	Q9NR68 homo sapien
11	11	5.3	149	2 Q25237	Q25237 lucilia cup
12	11	5.3	150	2 Q06784	Q06784 haematobia
13	11	5.3	175	2 Q6PLJ9	Q6PLJ9 squilla ora
14	11	5.3	175	2 Q9GN96	Q9GN96 chrysomya b
15	11	5.3	175	2 Q9GSU6	Q9GSU6 chrysomya b
16	11	5.3	175	2 Q9GSM0	Q9GSM0 chrysomya b

17	11	5.3	175	2 Q9GSM1	Q9GSM1 chrysomya b
18	11	5.3	175	2 Q9GSM4	Q9GSM4 chrysomya b
19	11	5.3	175	2 AAT09986	AAT09986 squilla o
20	11	5.3	176	2 Q8KSD7	Q8KSD7 mus musculu
21	11	5.3	182	2 Q6PLJ6	Q6PLJ6 penaeus jap
22	11	5.3	182	2 Q6PLJ8	Q6PLJ8 procambar
23	11	5.3	182	2 AAT09987	AAT09987 procambar
24	11	5.3	182	2 AAT09989	AAT09989 penaeus j
25	11	5.3	183	2 Q6PLJ7	Q6PLJ7 fennerope
26	11	5.3	183	2 AAT09988	AAT09988 mus musculu
27	11	5.3	200	2 Q924U6	Q924U6 mus musculu
28	10	4.8	50	2 Q9GTK6	Q9GTK6 culx quing
29	10	4.8	85	2 Q8MVL1	Q8MVL1 boltenia vi
30	10	4.8	116	2 Q7JMX5	Q7JMX5 helicoverpa
31	10	4.8	117	2 Q9PUF3	Q9PUF3 bothrops ja
32	10	4.8	125	2 Q804G0	Q804G0 sphecoidea
33	10	4.8	141	2 Q6GKZ6	Q6GKZ6 drosophila
34	10	4.8	142	2 Q8HYM3	Q8HYM3 felis silve
35	10	4.8	147	2 Q8HZD0	Q8HZD0 saguinus oe
36	10	4.8	149	2 Q6DTH8	Q6DTH8 hypophchalm
37	10	4.8	154	2 O18448	O18448 helicoverpa
38	10	4.8	155	2 Q9Y1K4	Q9Y1K4 anopheles g
39	10	4.8	159	2 Q8HZD1	Q8HZD1 macaca sp.
40	10	4.8	159	2 Q8HZD2	Q8HZD2 pongo pygma
41	10	4.8	159	2 Q8HZD3	Q8HZD3 gorilla gor
42	10	4.8	159	2 Q8HZD4	Q8HZD4 pan troglod
43	10	4.8	160	2 Q6XGZ1	Q6XGZ1 homo sapien
44	10	4.8	160	2 AAP70247	AAP70247 homo sapi
45	10	4.8	178	2 Q93594	Q93594 dicentrarch

ALIGNMENTS

RESULT 1  
Q9TV90 ID Q9TV90 PRELIMINARY; PRT; 73 AA.  
AC Q9TV90;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2004 (TREMBLrel. 26, Last annotation update)  
DE Plasminogen (Fragment).  
GN Name=PLG;  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99160468; PubMed=10051323;  
RA Caetano A.R., Pomp D., Murray J.D., Bowling A.T.;  
RT "Comparative mapping of 18 equine type I genes assigned by somatic  
cell hybrid analysis";  
RL Mamm. Genome 10:271-276(1999).  
DR EMBL; AF097581; AAD25984.1; -  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1  
FT NON\_TER 73  
SQ SEQUENCE 73 AA; 7826 MW; 6CCBBFBA93D07704 CRC64;

Alignment Scores:  
Pred. No.: 1-21e-30 Length: 73  
Score: 39.00 Matches: 39  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 18.84% Indels: 0

```

DB: 2 Gaps: 0
US-09-992-095B-53_COPY_1044_1664 (1-621) x Q9TV90 (1-73)
QY 451 GACAGTTCGCCAGGTGCACAGTGGAGTCTCTGTTTCTTCGAGAGGACAAATACATT 510
DB 35 AspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGluLysAspLysTyrIle 54
QY 511 TTACAGAGTCACTCTCTTGGGGTCTTGGGTGTGTGCACGCCCAATAAGCCTGGTGTGTC 567
DB 55 LeuGlnGlyValThrSerTrpGlyLeuGlyCysAlaArgProAsnLysProGlyVal 73

RESULT 2
Q6JDI3
ID Q6JDI3 PRELIMINARY; PRT; 54 AA.
AC Q6JDI3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RP PubMed=15233990;
RA Housley D.J.E., Ritzert E., Venta P.J.;
RT "Comparative radiation hybrid map of canine chromosome 1 incorporating
RT SNP and indel polymorphisms.";
RL Genomic 84:248-264(2004).
DR EMBL; AY514750; AAT44581.1; -.
DR InterPro; IPR001254; Peptidase_S1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydroxylase; Protease; Serine protease.
KW Hydroxylase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 54
FT NON_TER 54
SQ SEQUENCE 54 AA; 5865 MW; CEF792BED3F4281D CRC64;

Alignment Scores:
Pred. No.: 2,238-23 Length: 54
Score: 32.00 Matches: 32
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.46% Indels: 0
DB: 2 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x Q6JDI3 (1-54)
QY 436 TTGCGCGGAGGCACTGCACAGTTCGCCAGGTGCACAGTGGAGTCTCTGTTTCTTCGAG 495
DB 23 LeuAlaGlyGlyThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCysPheGlu 42
QY 496 AAGGACAAATACATTTCAGAGTCACTCTTCTGG 531
DB 43 LysAspLysTyrIleLeuGlnGlyValThrSerTrp 54

RESULT 3
Q6SLL2
ID Q6SLL2 PRELIMINARY; PRT; 176 AA.
AC Q6SLL2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Spee B., Penning L.C., Rothuizen J.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY455801; AAR19224.1; -.
DR GO; GO:0016301; F:Kinase activity; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Kinase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 176
FT NON_TER 176
SQ SEQUENCE 176 AA; 19624 MW; 6B5B70BF55FAE708 CRC64;

Alignment Scores:
Pred. No.: 7,89e-05 Length: 176
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.76% Indels: 0
DB: 2 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x Q6SLL2 (1-176)
QY 448 ACTGACAGTTCGCCAGGTGCACAGTGGAGTCTCTGTTTCTTCGAG 489
DB 157 ThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 170

RESULT 4
AAR19224
ID AAR19224 PRELIMINARY; PRT; 176 AA.
AC AAR19224;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Spee B., Penning L.C., Rothuizen J.;
RT "Differential gene expression of regenerative and fibrotic pathways in
RT canine hepatic portosystemic shunt and portal vein hypoplasia.";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY455801; AAR19224.1; -.
DR Kinase.
FT NON_TER 1
FT NON_TER 176
FT NON_TER 176
SQ SEQUENCE 176 AA; 19624 MW; 6B5B70BF55FAE708 CRC64;

Alignment Scores:
Pred. No.: 7,89e-05 Length: 176
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.76% Indels: 0
DB: 2 Gaps: 0

US-09-992-095B-53_COPY_1044_1664 (1-621) x AAR19224 (1-176)
QY 448 ACTGACAGTTCGCCAGGTGCACAGTGGAGTCTCTGTTTCTTCGAG 489
DB 157 ThrAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 170

```



RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
[7]  
FN SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6J; TISSUE=Tongue;  
RC ADACHI J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK019102; BAB31548.1; -  
DR EMBL; AK009451; BAB26297.1; -  
DR MEROPS; S01.257; -  
DR MGI; MGI:1929977; 2310015I08Rik.  
DR GO; GO:0005576; C:extracellular; IDA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON TER 1  
SQ SEQUENCE 87 AA; 9549 MW; B988D0CD62926EAA CRC64;  
Alignment Scores:  
Pred. No.: 0.00958 Length: 87  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 2 Gaps: 0  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x Q9CQ78 (1-87)  
QY 451 GACAGTTCGACGGTGACAGTGGAGTCTCTGGTTTC 489  
DB 34 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 46  
RESULT 6  
Q91VQ8 PRELIMINARY; PRT; 176 AA.  
AC Q91VQ8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 26, Last annotation update)  
DE BC010843 protein (Fragment).  
GN Name=BC010843  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=mix FVB/N;  
RC STRAIN=mix FVB/N;  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Haughey B.J., Ho J., Iacono J., et al. 1996.  
RA 'Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Basavant T.L., Scheetz T.E.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
[7]  
FN SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6J; TISSUE=Tongue;  
RC ADACHI J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK019102; BAB31548.1; -  
DR EMBL; AK009451; BAB26297.1; -  
DR MEROPS; S01.257; -  
DR MGI; MGI:1929977; 2310015I08Rik.  
DR GO; GO:0005576; C:extracellular; IDA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON TER 1  
SQ SEQUENCE 87 AA; 9549 MW; B988D0CD62926EAA CRC64;  
Alignment Scores:  
Pred. No.: 0.00958 Length: 87  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 2 Gaps: 0  
US-09-992-095B-53\_COPY\_1044\_1664 (1-621) x Q9CQ78 (1-87)  
QY 451 GACAGTTCGACGGTGACAGTGGAGTCTCTGGTTTC 489  
DB 34 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 46  
RESULT 6  
Q91VQ8 PRELIMINARY; PRT; 176 AA.  
AC Q91VQ8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 26, Last annotation update)  
DE BC010843 protein (Fragment).  
GN Name=BC010843  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=mix FVB/N;  
RC STRAIN=mix FVB/N;  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Haughey B.J., Ho J., Iacono J., et al. 1996.  
RA 'Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Basavant T.L., Scheetz T.E.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
[7]  
FN SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6J; TISSUE=Tongue;  
RC ADACHI J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.,  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK019102; BAB31548.1; -  
DR EMBL; AK009451; BAB26297.1; -  
DR MEROPS; S01.257; -  
DR MGI; MGI:1929977; 2310015I08Rik.  
DR GO; GO:0005576; C:extracellular; IDA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON TER 1  
SQ SEQUENCE 87 AA; 9549 MW; B988D0CD62926EAA CRC64;  
Alignment Scores:  
Pred. No.: 0.00958 Length: 87  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches

RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=mix FVB/N;  
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;  
RA Strausberg R.;  
RL submitted (JUL-2001) to the ENBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; BC010843; AAH10843.1; -;  
DR HSSP; P00760; IEZX.  
DR MEROPS; S01.087; -;  
DR MGD; MGI:2682935; BC010843.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1  
SQ SEQUENCE 176 AA; 19809 MW; F3B8C573C07E9717 CRC64;

Alignment Scores:  
Pred. No.: 0.000856 Length: 176  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.28% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q91VQ8 (1-176)

QY 451 GACAGTCCAGGGTGACAGTGAGGTCCTCTGTTGC 489  
|||||  
Db 114 AspSerCysGlnGlyAspSerGlyGlyProLeuValCys 126  
|||||

RESULT 7  
QY7M325 PRELIMINARY; PRT; 126 AA.  
AC QY7M325; (1-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
DE Chymotrypsin-like proteinase (EC 3.4.21.-) (Fragments).  
OS Sus scrofa domestica (domestic pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9825;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=86194934; PubMed=3634756;  
RA Vered M., Gertler A., Burshtein Y.;  
RL Int. J. Pept. Protein Res. 27:183-190(1986).  
DR PIR; A23473; A23473.  
DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR Pfam; PF00089; Trypsin; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1  
SQ SEQUENCE 126 AA; 12935 MW; E576321AEBF129CD CRC64;

Alignment Scores:  
Pred. No.: 0.00979 Length: 126  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.80% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q7M325 (1-126)

QY 454 AGTGGCCAGGGTGACAGTGAGGTCCTCTGTTGC 489  
|||||  
Db 95 SerCysGlnGlyAspSerGlyGlyProLeuValCys 106  
|||||

RESULT 8  
QYDC82 PRELIMINARY; PRT; 164 AA.  
AC QYDC82;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Mus musculus adult male spleen cDNA, RIKEN full-length enriched  
DE library, clone:0910001G08 product:chymotrypsin-like, full insert  
DE sequence.  
DE Name=Ctrl;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spleen;  
RX MEDLINE=99279253; PubMed=10349636;  
RA Carninci P., Hayashizaki Y.;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spleen;  
RX MEDLINE=21085660; PubMed=11217851;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spleen;  
RA The FANTOM Consortium;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spleen;  
RX MEDLINE=20499374; PubMed=11042159;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spleen;

RX MEDLINE=20530913; PubMed=11076861;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,  
RA "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multicapillary sequencer";  
RL Genome Res. 10:11757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Spleen;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
RA Muramatsu M., Hayashizaki Y.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK003074; BAB22549.1; -;  
DR HSSP; P00746; IFDP.  
DR MEROPS; S01.256; -;  
DR MGDI; MG1.88558; Ctrl.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR Hydrolyase; Protease; Serine protease.  
KW Hydrolyase; Protease; Serine protease.  
SQ SEQUENCE 164 AA; 17707 MW; 83791FD929ABEDD6 CRC64;  
Alignment Scores:  
Pred. No.: 0.00938 Length: 164  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.80% Indels: 0  
DB: 2 Gaps: 0  
US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q9DC82 (1-164)  
QY 454 AGTTGCGAGGTGACAGTGGAGTCTCTCTGTTGC 489  
Db 109 SerCysGlnGlyAepSerGlyGlyProLeuValCys 120  
RESULT 9  
TRY3\_LUCCU STANDARD; PRT; 165 AA.  
AC P35043;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Trypsin alpha-3 (EC 3.4.21.4) (Fragment).  
OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
OC Calliphoridae; Lucilia.  
OC NCBI\_TaxID=7375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95202116; PubMed=7894748;  
RA Casu R.E., Jarney J.M., Elvin C.M., Eisemann C.H.;

RT "Isolation of a trypsin-like serine protease gene family from the  
RT sheep blowfly *Lucilia cuprina*.";  
RL Insect Mol. Biol. 3:159-170(1994).  
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Arg|-Xaa, Lys|-Xaa.  
CC -|- SUBCELLULAR LOCATION: Extracellular.  
CC -|- SIMILARITY: Belongs to peptidase family S1.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; L15632; AAA65931.1; -;  
CC HSSP; P00763; ADPO.  
DR MEROPS; S01.112; -;  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; PARTIAL.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR Hydrolyase; Multigene family; Serine protease.  
KW Hydrolyase; Multigene family; Serine protease.  
FT NON\_TER 1 1  
FT ACT\_SITE 26 26 Charge relay system (By similarity).  
FT ACT\_SITE 119 119 Charge relay system (By similarity).  
FT DISULFID 89 106 By similarity.  
FT DISULFID 115 139 By similarity.  
FT SITE 113 113 Required for specificity (By similarity).  
SQ SEQUENCE 165 AA; 16569 MW; 26160B1AFF80F1CD CRC64;  
Alignment Scores:  
Pred. No.: 0.00937 Length: 165  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.80% Indels: 0  
DB: 1 Gaps: 0  
US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x TRY3\_LUCCU (1-165)  
QY 451 GACAGTTCAGGTGACAGTGGAGTCTCTCTGTT 486  
Db 113 AepSerCysGlnGlyAepSerGlyGlyProLeuVal 124  
RESULT 10  
Q9NR68 PRELIMINARY; PRT; 119 AA.  
AC Q9NR68;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Serine protease kallikrein/ovasin/neurosin type 3.  
DE Name=KLK8;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21206420; PubMed=11309326;  
RA Magklara A., Scorilas A., Katsaros D., Maassobrio M., Yousef G.M.,  
RA Fracchioli S., Danese S., Diamandis E.P.;  
RA "The human KLK8 (neurosin/ovasin) gene: identification of two novel  
RT splice variants and its prognostic value in ovarian cancer";  
RL Clin. Cancer Res. 7:806-811(2001).  
DR EMBL; AF251125; AAF79144.1; -;  
DR HSSP; P00760; 1EZX.

DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 SQ SEQUENCE 119 AA; 12718 MW; 2FD8164DF1641FFF CRC64;

Alignment Scores:  
 Pred. No.: 0.107 Length: 119  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.31% Indels: 0  
 DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q9NR68 (1-119)

QY 457 TGCCAGGTGACAGTGAGGTCCTCTGTTGC 489  
 DB 67 CysGlnGlyAaspSerGlyGlyProLeuValCys 77

RESULT 11  
 Q25237  
 ID Q25237 PRELIMINARY; PRT; 149 AA.  
 AC Q25237  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Serine proteinase (Fragment).  
 OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 OC Calliphoridae; Lucilia.  
 OX NCBI\_TaxID=7375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN:Long Pocket Laboratory culture; TISSUE=Total organism;  
 RX MEDLINE=95079073; PubMed=7987520;  
 RA Elvin C.M., Vuocolo T., Smith W.J., Eismann C.H., Riddles P.W.;  
 RT "An estimate of the number of serine protease genes expressed in sheep  
 blowfly larvae (Lucilia cuprina).";  
 RL Insect Mol. Biol. 3:105-115(1994).  
 CC -1- SIMILARITY: Belongs to peptidase family S1.  
 DR EMBL; U07693; AAL17385.1; -.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 149  
 SQ SEQUENCE 149 AA; 15212 MW; 95953C3945C317CD CRC64;

Alignment Scores:  
 Pred. No.: 0.103 Length: 149  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.31% Indels: 0  
 DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q25237 (1-149)

QY 451 GACAGTTCAGGTCAGCTGACAGTGAGGTCCTCTG 483  
 DB 139 AspSerCysGlnGlyAaspSerGlyGlyProLeu 149

## RESULT 12

Q06784  
 ID Q06784 PRELIMINARY; PRT; 150 AA.  
 AC Q06784  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Serine protease (Fragment).  
 OS Haemaphysalis irritans (Horn fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
 OC Muscidae; Haematobia.  
 OX NCBI\_TaxID=7368;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=93341451; PubMed=8341258;  
 RX Elvin C.M., Whan V.A., Riddles P.W.;  
 RA "A family of serine protease genes expressed in adult buffalo fly  
 (Haemaphysalis irritans exigua).";  
 RT Mol. Gen. Genet. 240:132-139(1993).  
 RL -1- SIMILARITY: Belongs to peptidase family S1.  
 CC EMBL; Z22567; CAA80289.1; -.  
 DR PIR; S35208; S35208.  
 DR MEROPS; S01.110; -.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR009003; Pept\_Ser\_Cys.  
 DR Pfam; PF00089; Trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 KW Hydrolase; Protease; Serine protease.  
 FT NON\_TER 150  
 SQ SEQUENCE 150 AA; 16047 MW; 290D3054697FF37F CRC64;

Alignment Scores:  
 Pred. No.: 0.103 Length: 150  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.31% Indels: 0  
 DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q06784 (1-150)

QY 451 GACAGTTCAGGTCAGCTGACAGTGAGGTCCTCTG 483  
 DB 140 AspSerCysGlnGlyAaspSerGlyGlyProLeu 150

## RESULT 13

Q06PLJ9  
 ID Q06PLJ9 PRELIMINARY; PRT; 175 AA.  
 AC Q06PLJ9  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Trypsin (Fragment).  
 OS Squilla oratoria.  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Hoplocarida; Stomatopoda; Squillidae; Squilla.  
 OX NCBI\_TaxID=274635;

Alignment Scores:  
 Pred. No.: 0.103 Length: 175  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.31% Indels: 0  
 DB: 2 Gaps: 0

RN SEQUENCE FROM N.A.  
RP Wang N., Jiang G., Li N., Li L., Zhang J., Ma X., Chen Q.;  
RA Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
RL -1- SIMILARITY: Belongs to peptidase family S1.  
CC EMBL; AY596941; AAT0986.1; -.  
DR InterPro: IPR001254; Peptidase\_S1.  
DR InterPro: IPR001314; Peptidase\_S1A.  
DR InterPro: IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1 1  
FT NON\_TER 175 175  
SQ SEQUENCE 175 AA; 18536 MW; 671EEABD5D1FF20A CRC64;

Alignment Scores:  
Pred. No.: 0.101 Length: 175  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q6PLJ9 (1-175)

QY 451 GACAGTTCACAGGTGACAGTGGAGTCTCTG 483  
DB 152 AppSerCysGInglyAspSerGlyGlyProLeu 162  
|||||

## RESULT 14

Q9GN96 PRELIMINARY; PRT; 175 AA.  
ID Q9GN96  
AC Q9GN96; (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 27, Last annotation update)  
DE Serine protease K4.1/F1R2 (Serine protease K6.1/F1R1) (Serine protease K6.2/F1R2) (Fragment).  
DE K6.2/F1R2 (Fragment).  
OS Chrysomya bezziana (Old world screwworm).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
OC Calliphoridae; Chrysomya.  
OX NCBI\_TaxID=69364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21411890; PubMed=11520682;  
RA Muharsini S., Dalrymple B., Vuocolo T., Hamilton S., Willadsen P.,  
RA Wjffels G.;  
RT "Biochemical and molecular characterization of serine proteases from larvae of Chrysomya bezziana, the Old World Screwworm fly.";  
RT Insect Biochem. Mol. Biol. 31:1029-1040(2001).  
RL Insect Biochem. Mol. Biol. 31:1029-1040(2001).  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; AF302489; AAG30250.1; -.  
DR EMBL; AF302487; AAG30248.1; -.  
DR EMBL; AF302480; AAG30241.1; -.  
DR HSP; P00761; 1EPT.  
DR MEROPS; S01.110; -.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro: IPR001254; Peptidase\_S1.  
DR InterPro: IPR001314; Peptidase\_S1A.  
DR InterPro: IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1 1  
FT NON\_TER 175 175  
SQ SEQUENCE 175 AA; 18200 MW; 0AA9F30746F8FD3D CRC64;

## Alignment Scores:

Pred. No.: 0.101 Length: 175  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q9GSL6 (1-175)

QY 451 GACAGTTCACAGGTGACAGTGGAGTCTCTG 483  
|||||

DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1 1  
FT NON\_TER 175 175  
SQ SEQUENCE 175 AA; 18166 MW; 9DF0C459D83DC173 CRC64;

## Alignment Scores:

Pred. No.: 0.101 Length: 175  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q9GN96 (1-175)

QY 451 GACAGTTCACAGGTGACAGTGGAGTCTCTG 483  
DB 165 AppSerCysGInglyAspSerGlyGlyProLeu 175  
|||||

## RESULT 15

Q9GSL6 PRELIMINARY; PRT; 175 AA.  
ID Q9GSL6  
AC Q9GSL6; (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 26, Last annotation update)  
DE Serine protease K16/F1R2 (Fragment).  
DE Chrysomya bezziana (Old world screwworm).  
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
OC Calliphoridae; Chrysomya.  
OX NCBI\_TaxID=69364;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21411890; PubMed=11520682;  
RA Muharsini S., Dalrymple B., Vuocolo T., Hamilton S., Willadsen P.,  
RA Wjffels G.;  
RT "Biochemical and molecular characterization of serine proteases from larvae of Chrysomya bezziana, the Old World Screwworm fly.";  
RT Insect Biochem. Mol. Biol. 31:1029-1040(2001).  
RL Insect Biochem. Mol. Biol. 31:1029-1040(2001).  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
DR EMBL; AF302486; AAG30247.1; -.  
DR MEROPS; S01.110; -.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro: IPR001254; Peptidase\_S1.  
DR InterPro: IPR001314; Peptidase\_S1A.  
DR InterPro: IPR009003; Pept\_Ser\_Cys.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; TRYP\_SPC; 1.  
DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
KW Hydrolase; Protease; Serine protease.  
FT NON\_TER 1 1  
FT NON\_TER 175 175  
SQ SEQUENCE 175 AA; 18200 MW; 0AA9F30746F8FD3D CRC64;

## Alignment Scores:

Pred. No.: 0.101 Length: 175  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.31% Indels: 0  
DB: 2 Gaps: 0

US-09-992-095b-53\_COPY\_1044\_1664 (1-621) x Q9GSL6 (1-175)

QY 451 GACAGTTCACAGGTGACAGTGGAGTCTCTG 483  
|||||

Db 165 AspSerCysGlnGlyAspSerGlyGlyProLeu 175

Search completed: October 27, 2004, 09:11:41  
Job time : 118.5 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 27, 2004, 06:00:25 ; Search time 127 Seconds  
(without alignments)  
1050.486 Million cell updates/sec

Title: US-09-992-095b-53\_COPY\_1044\_1664  
Perfect score: 621  
Sequence: 1 argcaactctgtggaggcac.....aggagtgatgagaataat 621

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 672930 seqs, 107416849 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1279806

Minimum DB seq length: 0  
Maximum DB seq length: 621

Post-processing: Listing first 45 summaries

Database : Pending Patents NA New.\*  
1: /cgn2\_6/prodata/2/pna/PCT\_NEW\_COMB.seq.\*  
2: /cgn2\_6/prodata/2/pna/US06\_NEW\_COMB.seq.\*  
3: /cgn2\_6/prodata/2/pna/US07\_NEW\_COMB.seq.\*  
4: /cgn2\_6/prodata/2/pna/US08\_NEW\_COMB.seq.\*  
5: /cgn2\_6/prodata/2/pna/US09\_NEW\_COMB.seq.\*  
6: /cgn2\_6/prodata/2/pna/US10\_NEW\_COMB.seq.\*  
7: /cgn2\_6/prodata/2/pna/US11\_NEW\_COMB.seq.\*  
8: /cgn2\_6/prodata/2/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393	63.3	586	US-10-948-737-12785	Sequence 12785, A
2	382	61.5	489	US-10-948-737-13963	Sequence 13963, A
3	275	44.3	559	US-10-948-737-6644	Sequence 6644, Ap
4	94	15.1	600	US-10-956-157-10443	Sequence 10443, A
5	94	15.1	621	US-10-956-157-5208	Sequence 5208, Ap
6	25	4.0	25	US-10-956-157-18279	Sequence 18279, A
7	25	4.0	25	US-10-956-157-18280	Sequence 18280, A
8	25	4.0	25	US-10-956-157-18281	Sequence 18281, A
9	25	4.0	25	US-10-956-157-18285	Sequence 18285, A
10	25	4.0	25	US-10-956-157-18287	Sequence 18287, A
11	25	4.0	25	US-10-956-157-18291	Sequence 18291, A
12	25	4.0	25	US-10-956-157-18293	Sequence 18293, A
13	25	4.0	25	US-10-956-157-18294	Sequence 18294, A
14	25	4.0	25	US-10-956-157-18295	Sequence 18295, A
15	25	4.0	25	US-10-956-157-18296	Sequence 18296, A
16	25	4.0	25	US-10-956-157-159740	Sequence 159740, A
17	25	4.0	25	US-10-956-157-161962	Sequence 161962, A
18	25	4.0	25	US-10-956-157-162626	Sequence 162626, A
19	25	4.0	25	US-10-956-157-165143	Sequence 165143, A
20	25	4.0	25	US-10-956-157-167224	Sequence 167224, A
21	25	4.0	25	US-10-956-157-169162	Sequence 169162, A
22	25	4.0	25	US-10-956-157-170509	Sequence 170509, A
23	25	4.0	25	US-10-956-157-177180	Sequence 177180, A
24	25	4.0	25	US-10-956-157-190611	Sequence 190611, A
25	25	4.0	25	US-10-956-157-193277	Sequence 193277, A

26	25	4.0	25	6	US-10-956-157-205399	Sequence 205399, A
27	25	4.0	25	6	US-10-956-157-206729	Sequence 206729, A
28	25	4.0	25	6	US-10-956-157-212147	Sequence 212147, A
29	25	4.0	25	6	US-10-956-157-215631	Sequence 215631, A
30	25	4.0	25	6	US-10-956-157-216185	Sequence 216185, A
31	25	4.0	25	6	US-10-956-157-220676	Sequence 220676, A
32	25	4.0	25	6	US-10-956-157-229677	Sequence 229677, A
33	25	4.0	25	6	US-10-956-157-252159	Sequence 252159, A
34	25	4.0	25	6	US-10-956-157-256433	Sequence 256433, A
35	25	4.0	25	6	US-10-956-157-287208	Sequence 287208, A
36	25	4.0	25	6	US-10-956-157-294896	Sequence 294896, A
37	25	4.0	25	6	US-10-956-157-301354	Sequence 301354, A
38	25	4.0	25	6	US-10-956-157-301706	Sequence 301706, A
39	25	4.0	25	6	US-10-956-157-302834	Sequence 302834, A
40	25	4.0	25	6	US-10-956-157-306106	Sequence 306106, A
41	24	3.9	25	6	US-10-956-157-204559	Sequence 204559, A
42	23	3.7	25	6	US-10-956-157-209900	Sequence 209900, A
43	23	3.7	25	6	US-10-956-157-219147	Sequence 219147, A
44	23	3.7	25	6	US-10-956-157-264455	Sequence 264455, A
45	21	3.4	25	6	US-10-956-157-315519	Sequence 315519, A

ALIGNMENTS

RESULT 1  
US-10-948-737-12785  
; Sequence 12785, Application US/10948737  
; GENERAL INFORMATION:  
; APPLICANT: CHAN, VIVIEN W.  
; APPLICANT: ESCOBEDO, JAIME  
; APPLICANT: GARCIA, PABLO DOMINGUEZ  
; APPLICANT: HANSEN, RHONDA  
; APPLICANT: KAUFMANN, JOERG  
; APPLICANT: KENNEDY, GIULIA C.  
; APPLICANT: LAMSON, GEORGE  
; APPLICANT: MOLER, EDWARD J.  
; APPLICANT: RANDAZZO, FILIPPO  
; APPLICANT: REINHARD, CHRISTOPH  
; APPLICANT: SUDDUTH-KLINGER, JULIE  
; TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED  
; FILE REFERENCE: 2300-21987  
; CURRENT APPLICATION NUMBER: US/10/948,737  
; PRIOR FILING DATE: 2004-09-22  
; PRIOR APPLICATION NUMBER: 10/616,900  
; PRIOR FILING DATE: 2003-07-09  
; PRIOR APPLICATION NUMBER: 09/872,850  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/208,871  
; PRIOR FILING DATE: 2000-06-02  
; PRIOR APPLICATION NUMBER: 10/081,519  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 60/270,959  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 10/310,673  
; PRIOR FILING DATE: 2003-01-08  
; PRIOR APPLICATION NUMBER: 60/336,613  
; PRIOR FILING DATE: 2001-12-04  
; PRIOR APPLICATION NUMBER: US03/00657  
; PRIOR FILING DATE: 2003-01-08  
; PRIOR APPLICATION NUMBER: 60/345,637  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: 10/081,124  
; PRIOR FILING DATE: 2002-02-21  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 13996  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12785  
; LENGTH: 586  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-948-737-12785







```
RESULT 7
US-10-956-157-18280
; Sequence 18280, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18280
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18280
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 553 AATAAGCCTGGTGTCTATGTTTCGTG 577
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 AATAAGCCTGGTGTCTATGTTTCGTG 25

RESULT 8
US-10-956-157-18281
; Sequence 18281, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18281
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18281
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 552 CAATAAGCCTGGTGTCTATGTTTCGT 576
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CAATAAGCCTGGTGTCTATGTTTCGT 25

RESULT 9
US-10-956-157-18285
; Sequence 18285, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18285
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18285
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 554 TGTCTATGTTTCGTGTTTCAAGGTTT 588
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TGTCTATGTTTCGTGTTTCAAGGTTT 25

RESULT 10
US-10-956-157-18287
; Sequence 18287, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18287
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18287
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 595 TGGATTGAGGAGTGATGAGAAATA 619
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TGGATTGAGGAGTGATGAGAAATA 25

RESULT 11
US-10-956-157-18291
; Sequence 18291, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18291
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18291
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 564 TGTCTATGTTTCGTGTTTCAAGGTTT 588
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TGTCTATGTTTCGTGTTTCAAGGTTT 25

RESULT 12
US-10-956-157-18293
; Sequence 18293, Application US/10956157
; GENERAL INFORMATION:
```

```
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18285
```

```
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 353 TCCTGTGATTGAGATAAAGTGTG 377
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TCCTGTGATTGAGATAAAGTGTG 25
```

```
RESULT 10
US-10-956-157-18287
; Sequence 18287, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18287
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18287
```

```
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 595 TGGATTGAGGAGTGATGAGAAATA 619
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TGGATTGAGGAGTGATGAGAAATA 25
```

```
RESULT 11
US-10-956-157-18291
; Sequence 18291, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18291
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18291
```

```
Query Match 4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 564 TGTCTATGTTTCGTGTTTCAAGGTTT 588
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 TGTCTATGTTTCGTGTTTCAAGGTTT 25
```

```
RESULT 12
US-10-956-157-18293
; Sequence 18293, Application US/10956157
; GENERAL INFORMATION:
```

```

; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18293
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18293

Query Match      4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 AATCGCTATGAGTTTCTGAATGAA 403
Db 1 AATCGCTATGAGTTTCTGAATGAA 25

RESULT 13
US-10-956-157-18294
; Sequence 18294, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18294
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18294

Query Match      4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 GCTATGAGTTTCTGAATGGAAGAGT 407
Db 1 GCTATGAGTTTCTGAATGGAAGAGT 25

RESULT 14
US-10-956-157-18295
; Sequence 18295, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18295
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18295

Query Match      4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 383 GCTATGAGTTTCTGAATGGAAGAGT 407
Db 1 GCTATGAGTTTCTGAATGGAAGAGT 25

RESULT 15
US-10-956-157-18296
; Sequence 18296, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18296
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18296

Query Match      4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 ATCGCTATGAGTTTCTGAATGGAAG 404
Db 1 ATCGCTATGAGTTTCTGAATGGAAG 25

Search completed: October 27, 2004, 08:37:11
Job time : 128 secs
```

```

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 ATCGCTATGAGTTTCTGAATGGAAG 404
Db 1 ATCGCTATGAGTTTCTGAATGGAAG 25

RESULT 15
US-10-956-157-18296
; Sequence 18296, Application US/10956157
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18296
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Probe Sequence
US-10-956-157-18296

Query Match      4.0%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 CAATCGCTATGAGTTTCTGAATGGA 402
Db 1 CAATCGCTATGAGTTTCTGAATGGA 25

Search completed: October 27, 2004, 08:37:11
Job time : 128 secs
```

**This Page Blank (uspto)**